Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems

by

Christopher V. Rao

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Chemical Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

2000

© Copyright by Christopher V. Rao 2000
All Rights Reserved

To my homies... for keepin' it real

Contents

List of Tables

List of Figures

Acknowledgments

I want first to acknowledge my thesis advisor, James B. Rawlings. His intellect and character provided an exceptional environment for learning. He always kept me focused on the big picture, yet still encouraged me to pursue all of my silly and crackpot ideas. His thoughtful and poignant questions, true to the Socratic tradition, were invaluable in my development. I am eternally grateful for all that he has done for me.

I have been fortunate also to collaborate with two genuine scholars: Stephen J. Wright and David Q. Mayne. During my stay at Argonne, Steve generously shared his time with me and patiently answered my naive questions. He provided also valuable guidance early in my graduate studies. David's vast knowledge of control and optimization, superb intellect, and distinguished career were a constant source of inspiration. His ideas and helpful suggestions were indispensable. He always kept me focused on the details, constantly challenging all of my silly and crackpot ideas. I wish also to thank my thesis committee for the time and effort spent evaluating this work.

To say my education and dissertation are a product solely of the classes and research would do horrible injustice to the truth. Far and away the most memorable and worthwhile experiences of my graduate education were the times spent with the eclectic personalities that comprise the Rawlings research group. In addition to providing the core of my education, they profoundly influenced who I am today.

I would like to begin by acknowledging Tony, Ken, and John Eaton. I want to thank Tony for teaching me the ropes and always providing lively debate. Ken, even though he refused to talk to me when he was graduate student, was a source of wisdom once he graduated. I want to thank John Eaton for answering numerous computer questions. I want also to acknowledge the young members of the group: Jenny, Matt, and Brian. Jenny was a gracious and pleasant officemate, and Matt was alway able to put his unique spin on things. Brian, I regret we did not have a chance to spend more time together.

The research group would not have been so vibrant had it not been for the constant presence of visitors from lands far away. I want to thank Olav for teaching me that, indeed, God is in the details and that the "Black Label" is truly sublime. The symbiotic Findeisen twins, Rolf and Peter, were always there to help me with whatever little problem I had and to discuss my research problems with.

I want next to acknowledge Rahul. Your demeanor was singularly pleasing, and I treasure the five years we spent together. Your wisdom on the ways and ironies of love were greatly appreciated, even though I failed to heed them.

The 1111 Gilson household is a constant source of fond memories: ice hockey, roller blading, Christmas, the flooding basement, the fork... Pierre, in addition to patiently teaching me MPC, was my Texas partner in crime. I will forever reminisce about our expeditions to 6th Street and the late nights at Trudy's. Rock and John, I am truly honored to call you my friends. Both of you taught me so much about so many different things. Rock, you were alway a constant source of exciting conversation and debate, and John, I doubt you realize how sorely you are missed.

Scott and Daniel, I cannot describe in few sentences what your friendship and love has meant to me. Scott, in addition to all else, you taught me what character is and how one should live the good life. Together we learned the indignities a person will suffer just for cheap rent. Daniel, all I can say is that I cannot imagine it without you and that one day I hope the "curse" is finally broken.

I conclude by acknowledging Theresa and Andrea for making this odyssey both exciting and memorable - you were the highs and lows.

CHRISTOPHER V. RAO

University of Wisconsin–Madison
February 2000

Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems

Christopher V. Rao
Under the supervision of Professor James B. Rawlings
At the University of Wisconsin–Madison

The rational design of process monitoring and control systems requires the solution of dynamic programs. With a few notable exceptions, dynamic programs are difficult, in not impossible, to solve. The difficulty arises in what Bellman called the "curse of dimensionality": the computational complexity scales exponentially in the problem dimensions. One approximate strategy that circumvents the computational difficulties associated with dynamic programming while still retaining many desirable properties is the moving horizon approximation. Moving horizon approximations are optimization based strategies that approximate the dynamic program with a series of open-loop optimal control problems. Unlike other strategies, moving horizon approximations can handle explicitly nonlinear differential algebraic equations and inequality constraints. In this dissertation, we investigate the moving horizon approximation for the constrained process monitoring (moving horizon estimation) and control (model predictive control) of nonlinear discrete-time systems. A framework is proposed for analyzing the stability properties of the moving horizon approximation. This framework allows us to derive sufficient conditions for stability and propose practical algorithms for online implementation.

In addition to the theoretical results, practical issues regarding constraints, computation, and robustness are studied. We discuss issues regarding inequality constraints in process monitoring. By incorporating prior knowledge in the form of inequality constraints, one can significantly improve the quality of state estimates for certain problems. We demonstrate how inequality constraints provide a flexible tool for complementing process knowledge and a strategy also for model simplification. For control, techniques are developed for handling inequality constraints active at steady state, a case that has not been treated in previous model predictive control theory

Computational issues are addressed. Stable suboptimal algorithms for constrained estimation and control are proposed that do not require an optimal solution: rather, a feasible solution suffices. Issues related to formulating model predictive control as a linear program are discussed. A computationally efficient interior point algorithm is developed for the model predictive control of large process systems. The cost of this approach is linear in the horizon length, compared with cubic growth for a naive approach. We also investigate strategies for further decomposing the problem structure in sheet and film forming processes.

The issue of output feedback and robustness are addressed by formulating MPC as a dynamic game. The game formulation allows us to obtain a separation for output feedback and prove that the closed-loop system has finite l_2-gain. Furthermore, the added cost associated with formulating MPC as a dynamic game is negligible; the resulting problem is a quadratic program, though the optimization problem is no longer sparse. These results are extremely conservative, however, and limitations of the proposed strategy are discussed.

Chapter 1

Introduction

The rational design of process monitoring and control strategies requires the solution of an optimal control problem. By rational design, we mean design by (sub)-optimal satisfaction of some specified objective and constraints. The characteristic features of these design problems are that the decisions are made in stages and, more importantly, with limited knowledge. Under these circumstances, dynamic programming provides the natural solution technique for these classes of problems. Without going into any detail (c.f. Bertsekas (1995*a*, 1995*b*)), dynamic programming recursively decomposes the overall problem by ranking decisions on the present cost and future (past) expected cost, assuming optimal decisions are made for the subsequent (previous) stages. What distinguishes dynamic programming from classical methods is that it is constructive. Unlike classical control theory rooted in operator theory and complex analysis, which provides only quantitative tests such as stability, dynamic programming generates sufficient conditions for optimal recursive state estimators and feedback policies. Since the 1950's, control engineers have recognized the potential of dynamic programming. The string of applications is impressive, in particular linear quadratic control (LQG) and Kalman filtering. Even strategies once thought to transcend optimal control and dynamic programming such as $\mathcal{H}_\infty$ control have, in their most general forms, been recast in the optimal control framework (c.f. (Başar and Bernhard 1995)). While it is easy to extol virtues of dynamic programming, there is one significant caveat: with a few notable exceptions, dynamic programs are difficult, if not impossible, to solve. The difficulties arise in what Bellman called the "curse of dimensionality": the computational complexity scales exponentially in the problem dimensions, and, as a result, limits the size of problems one can solve with modest computational recourses. Rather than despair, control engineers have introduced a host of strategies that allows one to approximately solve dynamic programs while still retaining many desirable properties. One generic and powerful strategy is the moving horizon approximation.

Moving horizon approximations arise in both process monitoring (moving horizon estimation) and control (model predictive control). The key idea behind the moving horizon approximation is to reformulate the dynamic program as a sequence of finite-horizon, open-loop optimal control problems, thereby restricting the algorithm's attention to the encountered sequence of states rather than the entire state space. The payoff is that a potentially intractable dynamic program is replaced with a sequence of computationally tractable open-loop optimal control problems. Whereas it is difficult, if not impossible, to solve dynamic programs, there exists a host of numerical strategies for solving finite-horizon open-loop optimal control problems (c.f. (Polak 1997)). The tradeoff is that the optimal value function is lost. However, constructing the value function is what complicates the solution of dynamic programs. In process monitoring, the optimal value function, or arrival cost, generates the conditional probability density function necessary for constructing recursive state estimators. If one considers only batch estimation problems, then this limitation is not problematic. However, if one wants to process information continuously or make decisions based on inferences, then a representation of the value function is necessary. In control, the value function, or cost to go, generates the optimal feedback policy.

Without the value function the feedback aspect of the solution is lost: instead of optimizing over policies, fixed control decisions are made. The implications of an open-loop strategy become apparent if we consider the issue of robustness and stochastic control. Of course, model predictive control is a feedback policy, but it is not an optimal feedback policy. An open question is how far model predictive control is from an optimal feedback policy. Current theory states that model predictive control performs no worse than pure open-loop control (Bertsekas 1972). Nevertheless, given our current repertoire of numerical algorithms, moving horizon approximations are practical methods for generating nearly optimal policies for problems involving inequality constraints and nonlinear dynamics.

Controlling nonlinear dynamics is undoubtedly a challenging and important problem, garnering tremendous interest from control theorists. However, one can legitimately argue that designing monitoring and control strategies that explicitly account for hard constraints on the controls and states has far greater impact in practice. Many applications require the satisfaction of constraints; for example, valves saturate and the controller needs to maintain the state variables, such as velocity, temperature, concentration, or pressure, within certain limits to operate the process safely. Furthermore, operation at constraints is so common that it may be regarded as the rule rather than the exception in chemical process operations. As a result model predictive control (also known as receding horizon control in the automatic control vernacular) with linear models is popular in the chemical process industries. There are over two thousand documented commercial applications, and nearly all refineries implement some form of model predictive control, the most popular being DMC and IDCOM. With the clarity gained by hindsight, the *raison d'etre* for model predictive control is the ability to handle constraints on the control and state variables. Though less mature than model predictive control, constraints motivate also the use of moving horizon estimation. While documented applications of moving horizon estimation and model predictive control with *nonlinear models* are gradually beginning to appear (c.f. (Russo and Young 1999) and (Qin and Badgwell 1998)) and show no signs of abating, *hard constraints* will continue to motivate most implementations of model predictive control and moving horizon estimation.

1.0.1 Process Monitoring and Moving Horizon State Estimation

Monitoring the dynamic behavior of a chemical process necessitates continual inferences from the available measurements about the evolving state of the process. To handle the ever growing influx of data, one needs some form of data compression to keep the problem size manageable. If one tackles the problem using probability theory, then one achieves compression through the use of a conditional probability density function generated by the solution of Kolmogorov's forward equation. If one instead solves the problem using deterministic least squares or game theory, then compression is achieved with the arrival cost obtained from the solution of a forward dynamic program. Both problems are mathematically equivalent even though they originate from different perspectives, and both are impossible to solve, analytically or numerically, in general, with the exception of linear unconstrained systems where one obtains a Kalman or $\mathcal{H}_\infty$ filter. Moving horizon estimation (MHE) bypasses the compression problem by considering only a fixed amount of data. The basic strategy is to estimate the state using a moving and fixed-size window of data. When a new measurement becomes available, the oldest measurement is removed from the data window and the newest measurement is added. The problem size of the estimation problem is bounded, therefore, by looking at only a subset of the available information. This strategy is obviously fraught with peril (otherwise, why fuss with compression). The reason is simple: unless judiciously constructed, fixed-horizon estimators may result in poor or, in the case of instability, catastrophic performance. However, we are able to construct stable moving horizon estimators that approach optimal performance. Furthermore, because MHE is formulated as an optimization problem, inequality constraints are a natural addition to the estimator.

As with control, the ability to handle inequality constraints explicitly is what makes moving horizon estimation attractive. One often has additional information about the process in the form of inequality constraints. For example, many process variables, such as concentrations, are positive. Also, in many practical situations we are able to provide hard bounds on the disturbance and state variables based on prior information, operating experience, and physical laws. In probabilistic terms, constraints may be used to model random variables with truncated or state-correlated probability densities. Constraints also allow the use of simplified or approximate models, where the inequality constraints complete the conservation laws of interest.

1.0.2 Model Predictive Control

A standard problem in control is to design a feedback law that minimizes an objective over an infinite horizon. The optimal solution to this problem can be obtained in principle by solving the Hamilton-Jacobi-Bellman (HJB) equation (the dynamic program that arises in control). This often is a difficult task. One exception is when the system is linear, the objectives are quadratic, and there are no hard constraints on the inputs or states. In this case, the optimal cost function can be parameterized as a symmetric matrix, and the feedback law reduces to a linear quadratic regulator. When either of these conditions is violated, general procedures for solving the HJB equation do not exist. Model prediction control (MPC), or equivalently receding horizon control, is a constructive optimization based strategy that avoids solving the HJB equation by repetitively solving an open-loop optimal control problem instead. This strategy is also referred to by the oxymoron, open-loop feedback control (e.g. (Dreyfus 1962)). While lacking the power and elegance of a closed-loop controller obtained from dynamic programming, model predictive control is a practical strategy that exploits the availability of inexpensive high-performance desktop computers.

Many issues and temptations arise in the design and implementation of model predictive control. A prime issue of importance is stability: optimal controllers are not necessarily stable. Care must be taken to ensure stability. Other important issues include target calculations, disturbance modeling, infeasibility, and computation. Model predictive control also tempts the engineer with designs that may yield nonintuitive results and tuning difficulties, such as linear programming formulations.

1.0.3 Uncertainty and Robustness

One implements a control system in order to introduce regulatory feedback in a process. Feedback allows for design flexibility and adaptivity. It improves the process's ability to attenuate exogenous (downstream) variations and allows the process to recover autonomously from unforeseen events and disruptions. This robustness offers tremendous benefit as operating procedures and conditions inevitably evolve from the original design objectives. One cannot imagine a competitive alternative to feedback in accounting for uncertainty in designs. The power of dynamic programming is that it generates sufficient conditions for an optimal feedback policy. As mentioned already, moving horizon strategies lack the feedback character of dynamic programming solutions. The payoff is obviously that we now have something that we can solve. However, if we treat uncertainty explicitly in our formulation of model predictive control, then the tradeoff become immediately obvious.

There are two popular alternatives for making decision with incomplete knowledge: the stochastic (or Bayes) solution and the minimax solution (c.f. (Arrow 1951)). Although one is rooted in probability theory and the other in game theory, the tangible difference in their representations of the uncertainty, both provide constructive procedures for designing controllers *robust* to uncertainty. By not accounting for feedback, the performance and viability of model predictive control in stochastic or minimax optimal control problems are limited. For example, in minimax optimal control, existence conditions are far more conservative for open-loop control than for closed-loop control (i.e. dynamic programming). Uncertainty, acting as an adversary, can easily choose a disturbance that trumps the fixed control. Hence, for a solution to exist with open-loop control, the magnitude of the disturbances needs to be small. With a feedback policy, however, the controller can always compensate for any disturbance, because the dynamic programming solution implicitly accounts for all possible scenarios.

In practice these technical issues are less important, because model predictive control is a feedback policy. Even though it is not an optimal policy, the performance is quite good as evidenced by its popularity among process control engineers. The problems arise when one directly includes uncertainty in the control design. The motivation for doing so is to guarantee certain stability margins in the design. Stability margins are difficult to quantify for arbitrary nonlinear controllers, so often the only way to ensure a stability margin is to include the specification *a priori* in the design.

1.1 Dissertation Overview

In this dissertation we investigate moving horizon approximations for systems described by nonlinear difference equations of the form

$$\begin{aligned} x_{k+1} &= f(x_k, u_k, w_k), \\ y_k &= h(x_k) + v_k, \end{aligned}$$

where it is known that the state of system x_k, control u_k, disturbance w_k, and measurement noise v_k satisfy the following constraints

$$x_k \in \mathbb{X}, \;\; u_k \in \mathbb{U}, \;\; w_k \in \mathbb{W}, \;\; v_k \in \mathbb{V}.$$

Our goal is to approximate the dynamic programs that arise in regulation and estimation using open-loop optimal control. If we consider control, then the optimal regulator is obtained from the solution, assuming it exists, of the following backward dynamic program (i.e, the Hamilton-Jacobi-Bellman equation)

$$V(x,w) = \min_{u \in \mathbb{U}} \left\{ l(u,x) + V(f(x,u,w)) \;:\; f(x,u,w) \in \mathbb{X} \right\},$$

where the (stationary) feedback policy is given by

$$\mu(x,w) = \arg\min_{u \in \mathbb{U}} \left\{ l(u,x) + V(f(x,u,w)) \;:\; f(x,u,w) \in \mathbb{X} \right\}.$$

The moving horizon approximation of the optimal regulator is obtained by repetitively solving the following open-loop optimal control problem

$$\min_{\{u_k\}_{k=0}^{N-1} \in \mathbb{U}^N} \left\{ \sum_{k=0}^{N-1} l(u_k, x_k) + F(x_N) \;:\; \begin{array}{c} x_0 = x \\ x_{k+1} = f(x_k, u_k, w_k) \\ x_k \in \mathbb{X} \end{array} \right\}. \tag{1.1}$$

If we let $\{u_k^*(x, \{w_k\})\}_{k=0}^{N-1}$ denote the solution to (1.1), then we define model predictive control as the feedback policy

$$\mu(x,w) = u_0^*(x, \{w_k\}).$$

If we consider estimation, then the optimal (recursive) estimator is determined from the solution of the following forward dynamic program

$$\mathcal{Z}_{k+1}(x, u_k, y_k) = \min_{z \in \mathbb{X},\, w \in \mathbb{W}} \left\{ L(w,v) + \mathcal{Z}_k(z) \;:\; \begin{array}{c} x = f(z, u_k, w) \\ f(z, u_k, w) \in \mathbb{X} \\ y_k - h(z) \in \mathbb{V} \end{array} \right\},$$

where the optimal estimate is given by

$$\hat{x}_{k+1} = \arg\min_{z \in \mathbb{X}} \mathcal{Z}_k(z, u_k, y_k).$$

The moving horizon approximation of the optimal estimator is obtained by repetitively solving the following open-loop optimal control problem

$$\min_{\{x_k\}_{k=T-N}^{T-1},\, \{w_k\}_{k=T-N}^{T-1}} \left\{ \sum_{k=T-N}^{T-1} L(w_k, v_k) + \hat{\mathcal{Z}}_{T-N}(x_{T-N}) \ : \ \begin{array}{c} x_k \in \mathbb{X},\ u_k \in \mathbb{U} \\ x_{k+1} = f(x_k, u_k, w_k) \\ y_k - h(x_k) \in \mathbb{V} \end{array} \right\}. \tag{1.2}$$

If we let $\{x_k^*(\{u_k\}, \{y_k\})\}_{k=T-N}^{T}$ and $\{w_k^*(\{u_k\}, \{y_k\})\}_{k=T-N}^{T-1}$ denote the solution to (1.2), then the moving horizon estimate is given by

$$\hat{x}_T = x_T^*(\{u_k\}, \{y_k\}).$$

Most of this dissertation is directed towards deriving sufficient conditions for stability, though practical and computational issues are also discussed. A major theme is the notion of the cost to go $V(\cdot)$ and the arrival cost $\mathcal{Z}_k(\cdot)$. These concepts from dynamic programming provide the key result of this dissertation–a set of dual inequalities for the terminal penalty $F(\cdot)$ in control and the initial penalty $\hat{\mathcal{Z}}_k(\cdot)$ in estimation sufficient to guarantee the stability of the moving horizon approximation. We specialize these results to the case when the model is linear, the objectives are quadratic, and the inequality constraints are polyhedral convex sets. The dissertation concludes with a discussion of robustness and output feedback.

The dissertation is organized in three main sections. The first section discusses moving horizon estimation. The second section discusses model predictive control. The third section discusses some issues regarding robustness and output feedback. The dissertation concludes with a summary of the main results and a discussion of some open research problems.

Chapter 2

Moving Horizon Estimation

2.1 Introduction

At the heart of any monitoring or control strategy is an automated decision process. To make a rational decision, one judges the various alternatives based on supporting factual evidence and then chooses the decision that yields the greatest expected utility. Sometimes the evidence necessary for making a rational decision is directly observable, though in the process environment, more often than not, it is not directly observable and we need to infer the primary evidence from secondary evidence. However, coupled with empirical correlations and physical laws distilled from experience, secondary observations are often sufficient to infer the condition of the process environment and obtain the evidence necessary for any decision process.

The coupling of observation with physical insight is what characterizes inference and estimation. The problems of inference and estimation have had a long and distinguished history in the natural and social sciences, engaging such luminaries as Gauss, Legendre, Wiener, Kolmogorov, and Krein. For many problems the inference process appears simply as an observation process. Implicit in the inference, however, are probabilistic assumptions such as normality that allow one to assess confidence in the supporting evidence gleaned from the observations. Without some measure of confidence or uncertainty, we cannot meaningfully maximize utility, because we are unable to characterize our expectations. For some problems observations alone do not provide sufficient evidence for making rational decisions. Rather, we need to couple the observations with explicit physical insight in the form of a mathematical model. The typical form of the process model is either a state representation or an input-output (Volterra) representation. Often, though not always, state representations arise from physico-chemical insight, whereas input-output representations arise typically from empirical correlations. If we delve further, input-output representations couple directly the observation process with our empirical insight, because the insight derives from trends and correlations in the observations. Consequently, the inference process is coupled directly with the observation process. When one makes an inferences using a state representation, the observation process is divorced in most cases from our physical insight distilled in the model, though, of course, the physical insight developed originally through logical empiricism.

This chapter discusses the dynamic inference problem using a state representation, also referred to as a dynamic state estimation problem. Many control and monitoring systems are based on state-space models. The state is a natural construct in the fundamental modeling of chemical and biological processes, because it compactly summarizes the past information needed to understand the future behavior of the process. For example, temperature, pressure, and concentrations comprise the state of a single phase chemically reactive system. Whether full spatial or simple functional representations such as lumping are employed depends on the accuracy required. However, rarely is the state directly available from the process measurements, and the state typically needs to be inferred from secondary process measurements or a measurable subset of the state. For example, the average molecular weight of many polymer systems is inferred from viscosity measurements. Also, the concentration in a simple chemically reactive system may be inferred from the reactor temperature, a more easily measured state variable.

The importance of state estimation in engineering is well recognized, and the problem has attracted significant attention for more than fifty years. One can legitimately argue that the seminal result in state estimation is the work of Kalman (1960*b*, 1961). Most, if not all, practical strategies for state estimation are direct relatives of the Kalman filter. Consequently, the goal of this chapter is not to propose an alternative to Kalman filtering or even offer new insight into the problem of state estimation. Our far less ambitious goal, rather, is a problem that Kalman filtering does not address, the issue of constraints.

For a subset of problems, one possesses insights in addition to physical laws and empirical correlations in the form of inequality constraints on the state variables and process uncertainties. A simple example is that many state variables, such as temperature and concentration, are positive. As we discuss, many constraints arise from physical insights that are quantified in physical law, though not explicitly in the standard form of differential algebraic equations. The differential algebraic equations satisfy the constraints directly through physical laws and appropriate boundary conditions. In the inverse problem, however, while the physical laws are still present, often the goal is to reconstruct the boundary conditions from the observations. One obtains useful answers in most inference problems without regard to constraints. For a class of problems, as we demonstrate through examples, in order to obtain meaningful answers, the inference process needs to enforce compliance of the constraints. We focus on this problem.

Satisfying inequality constraints is the domain of mathematical programming. Consequently any inference process that incorporates constraints is necessarily formulated as a mathematical program. Our interest is in the dynamic estimation problem. Hence, our solution employs online optimization. Whereas one can view Kalman filtering, among many different alternatives, as an online optimization strategy, and our proposed solution reduces to Kalman filtering when we do not consider constraints, the constrained estimation problem cannot escape online optimization. While providing the ability to incorporate constraints, online optimization introduces practical difficulties related to data compression. Our proposed solution is moving horizon estimation (MHE). As we demonstrate, MHE bypasses the compression issue, albeit approximately, and provides, in our opinion, a practical and flexible strategy for constrained state estimation.

2.2 Literature Review

State estimation encompasses such disparate fields as engineering, statistics, mathematics, geology, and econometrics. While people have worked on estimation problems since the time of Galileo, it was not until the 1940's that estimation was first studied systematically. This work involving such legendary mathematicians as Kolmogorov, Krein, and Wiener was purely theoretical, though Wiener's work was motivated partially by an anti-aircraft fire-control problem. The major breakthrough in linear estimation theory from a practical standpoint was the work of Kalman (1960*b*) and Kalman and Bucy (1961). Unlike the Wiener filter, which requires the solution of an integral equation (the Wiener-Hopf equation), the Kalman filter requires only the iteration of an ordinary difference equation or integration of a differential equation. The solution of this problem, unlike the Wiener filter, is suitable for implementation online. What distinguished the Kalman filter is that it uses a state representation of the process, whereas the Wiener filter uses an input-output or signal covariance representation. Kalman filtering, therefore, "*is not a triumph of applied probability theory: the theory has only a slight inheritance from probability theory while it has become an important pillar of systems theory*" (Kalman 1994). Since the publication of Kalman's papers, there has been an explosion of research activity on Kalman filtering and state estimation. As Kalman (1994) humbly notes, over two hundred thousand cumulative papers, technical reports, and books have been written on Kalman filtering. Two excellent historical accounts of linear estimation are given by Sorenson (1970) and Kailath (1974).

Unlike the linear problem, there have been no comparable breakthroughs in nonlinear state estimation, even though a general theory for nonlinear systems was developed long ago (Stratonovich 1960, Kushner 1964, Kushner 1967), The problem is that these results are impractical for application–they requires either the solution of a partial integro-differential equation or a functional integral difference equation–and, consequently, they are of theoretical significance only. One commonly constructs a nonlinear state estimator, therefore, by linearizing the dynamic system along the estimated trajectory and then employing Kalman filtering. The result is the extended Kalman filter. Jazwinski (1970) provides an excellent discussion of extended Kalman filtering and its application. While the success of the extended Kalman filter in application is widely documented in the literature, there are no theoretical stability results available other than local results (Song and Grizzle 1995), non-divergence conditions (Safonov and Athans 1978), and results for parameter estimation in linear systems (Ljung 1979).

Another approach for designing nonlinear state estimators is to transform the nonlinear system to a linear system using a local coordinate transformation by output injection (Bestle and Zeitz 1983, Krener and Isidori 1983). The theory is rooted in differential geometry. The idea is analogous to feedback linearization, where the nonlinear system is transformed to an equivalent linear system by precompensating with feedback. Isidori (1989) provides a general discussion of differential geometric techniques in estimation and control. The strength of these approaches is that they yield stable observers. However, linearizing transformations may exist only locally and are often difficult to obtain for complex system. Furthermore, because general linearizing coordinate transformations require the solution of a set of partial differential equations, solutions exist for only a limited class of systems. Another approach is to use variable structure (or sliding mode) strategies (Slotine, Hedrick and Misawa 1987). These strategies are rarely amenable to process systems however. The reader is directed to Muske and Edgar (1996) for a survey nonlinear state estimation. There is also the related field of nonlinear observers. An excellent, though slightly dated, comparative study of nonlinear observers is given by Walcott, Corless and Żak (1987).

The success of model predictive control has motivated many researchers to investigate online optimization strategies for constrained and nonlinear state estimation. The connection between optimization and estimation dates back to Galileo in 1632, who sought to minimize various functions of the prediction error. The method of least squares developed by Gauss in 1795 is also an optimization based strategy for estimation. Even the Wiener filter was derived originally using variational calculus. However, the original derivation of the Kalman filter used orthogonal projection. Bryson and Frazier (1963) first showed the connection between Kalman filtering and optimization.

MHE is a practical strategy to handle the computational difficulties associated with optimization based estimation, and, as a consequence, many authors have explored different issues in MHE. The first application of MHE for *nonlinear* systems was the work of Jang, Joseph and Mukai (1986). Their strategy ignores disturbances and constraints and attempts to estimate only the initial state of the system. Thomas (1975) and Kwon, Bruckstein and Kailath (1983) discussed earlier moving horizon strategies for *unconstrained linear* systems. Limited memory and adaptive filters for *linear* systems are analogous to MHE, because only a fixed window of data is considered (see Jazwinski (1970) for a discussion of limited memory filters). Many process systems researchers extended the work of Jang and coworkers. Bequette and coworkers (1991, 1993) investigated moving horizon strategies for state estimation as a logical extension of model predictive control. Edgar and coworkers (1991, 1992) investigated moving horizon strategies for *nonlinear* data reconciliation. Biegler and coworkers (1991, 1996, 1997) investigated statistical and numerical issues related to optimization based *nonlinear* data reconciliation. Marquardt and coworkers (1996, 1998) discussed multi-scale strategies for MHE and the benefits of incorporating *constraints* in estimation. Findeisen (1997) investigated the stability of unconstrained linear MHE with filtering and smoothing updates. Bemporad, Mignone and Morari (1999) discussed the application of MHE to *hybrid* systems. Gesthuisen and Engell (1998) discussed the application of MHE to a pilot-scale polymerization reactor and Russo and Young (1999) discussed the application of MHE to an industrial polymerization process at the Exxon Chemical Company. Because MHE is formulated as an optimization problem, it is possible to handle explicitly inequality constraints. Robertson and Lee (1995, 1996, 1998) have investigated the probabilistic interpretation of constraints in estimation. Muske and Rawlings (1993, 1995) derived some preliminary conditions for the stability of state estimation with inequality constraints. Tyler and Morari (1996, 1997) demonstrated how constraints may result in instability for non-minimum phase systems.

In parallel to the research done in process systems, unconstrained MHE was investigated also by researchers in automatic control. Zimmer (1994) investigated an unconstrained MHE strategy similar to the approach of Jang et al. (1986) and derived conditions for stability using fixed-point theorems. Moraal and Grizzle (1995) also derived conditions for stability using fixed-point theorems. However, Moraal and Grizzle (1995) formulated the estimation problem as the solution of a set of algebraic equations. Michalska and Mayne (1995) investigated an unconstrained MHE strategy similar to the approach of Jang et al. (1986) and derived conditions for stability using Lyapunov arguments.

What distinguishes this work is that we provide a theory of MHE that is compatible with deterministic estimation theory, including Kalman filtering. A major focus is on reconciling constraints, particularly those on state variables, with estimation theory. As we demonstrate, state constraints alter implicitly the problem structure. The outline of the Chapter is as follows. We begin by introducing the constrained estimation problem and then show how moving horizon estimation arises when one considers online implementation. Our focus then shifts, and we discuss constraints, in particular the probabilistic interpretation of state constraints and the issue of causality. Using a series of examples of varied complexity, we illustrate the potential utility of incorporating constraints in the inference process. We conclude with a summary of our investigations.

2.3 Constrained State Estimation

At time T suppose our observations of the process comprise solely of a sequence of discrete measurements $\{y_0, y_1, \dots, y_{T-1}\}$. For simplicity we limit our discussion to the problem where all of the sensors provide measurements simultaneously, though we can extent the proposed strategy *mutatis mutandis* to incorporate multi-rate sensors. The objective at time T is to reconstruct the evolution of the state of the process $\{x(t); t \geq 0\}$ from the observations $\{y_0, y_1, \dots, y_{T-1}\}$.

We assume we can capture our physical insight of the process with a finite-dimensional[1] differential algebraic equation of the form

$$F(x(t), \dot{x}(t), u(t), w(t), t) = 0, \tag{2.1}$$

where $\dot{x}(\cdot)$ denotes the time derivative of the state $x(\cdot)$, $u(\cdot)$ denotes measurable exogenous disturbances, and $w(\cdot)$ denotes unmeasurable exogenous disturbances. The disturbance $w(\cdot)$ is typically modeled as a random process and may account also for modeling uncertainty. If we couple our physical insight of the process with the measurements, then we require a model of the process sensors. We relate the observations $y(t)$ to state $x(t)$ using a model of the form

$$y(t) = g(x(t), t) + v(t), \tag{2.2}$$

[1]Extensions to "infinite-dimensional" or distributed parameter systems are possible, though this problem is far more complex.

where measurement uncertainty is captured in the vector $v(t)$. One commonly assumes the vector $v(t)$ is a normally distributed random variable. We stress that the vector $y(t)$ in (2.2) denotes the actual observation, and the vector $v(t)$ denotes the error between the observation $y(t)$ and the predicted sensor reading $g(x(t),t)$.

With the exception of linear and trivial nonlinear process models, we need to discretize the differential algebraic equation (2.1) in order to perform any computation or analysis. At this stage of our discussion, the discretization is conceptual. Discretization is usually performed during optimization. Whether one employs a simultaneous strategy (c.f. Biegler (1997, 1998) and Bock and coworkers (1998)) or discretizes first using a DAE solver (c.f. Ascher and Petzold (1998)) is inconsequential to our discussion, though extremely important when one considers online implementation. We suppose hereafter that the differential algebraic equation (2.1) is discretized with a zero-order hold on the disturbances $u(\cdot)$ and $w(\cdot)$ yielding the difference equation

$$x_{k+1} = f_d(x_k, u_k, w_k, k), \tag{2.3}$$

where the integer k denotes the discrete-time index. A typical choice is $t = k\Delta T$, where ΔT denotes the sampling period. The subscripts on the vectors x, u, w, and v denote the value at the points of discretization (e.g. $x_k = x(k\Delta T)$). We assume also the points of discretization (e.g. $t_k = k\Delta t$) coincides with the measurement times. Rarely is the equation $f_d(\cdot)$ in (2.3) available in algebraic form. Instead, we view the function $f_d(\cdot)$ abstractly as the numerical solution of (2.2) with initial condition x_k. The difference equation (2.3) consequently does not include explicitly algebraic constraints, even though the corresponding differential equation (2.1) does.

When we couple physical insight with the process measurements, we need to introduce a measure of uncertainty. The model predictions rarely, if ever, coincide with the process measurements. We need somehow to distribute the errors between the model and sensor measurements. In other words, we need to reconcile our model with the process measurements. Reconciliation in our framework amounts to a tradeoff between the vectors w_k and v_k. One may interpret w_k as process disturbances or model uncertainty and the vector v_k as sensor noise. A natural framework to characterize uncertainty is probability theory, where we treat the vectors w_k and v_k as random variables. Our choice of the respective probability distributions provides the reconciliation. A common alternative to probability theory is game theory. In game theory one uses instead, though with often the same result, deterministic uncertainty descriptions of the vectors w_k and v_k (c.f. Başar and Bernhard (1995)). Another alternative was proposed recently by Binder, Blank, Dahmen and Marquardt (1999). Eschewing both probability and game theory, they view the reconciliation problem instead as the inversion of a compact operator, an ill-posed problem. The tradeoff in their framework is the degree of regularization.

When we formulate the state estimation problem from the perspective of probability theory, we typically model the evolution of the state as a discrete-time Markov process[2]. As we expect the process measurements are correlated with the state, the quantity of interest becomes the conditional probability density function of the state evolution $\{x_0, x_1, \ldots, x_T\}$ given the process measurements $\{y_0, y_1, \ldots, y_{T-1}\}$:

$$p(x_0, x_1, \ldots, x_T \mid y_0, y_1, \ldots, y_{T-1}). \tag{2.4}$$

The optimal estimate of the state at time k given the measurements

$$\{y_0, y_1, \ldots, y_{T-1}\},$$

which we denote by $\hat{x}_{k|T-1}$, is then a functional $L_T(\cdot)$ of conditional probability density function (2.4):

$$\{\hat{x}_{0|T-1}, \hat{x}_{1|T-1}, \ldots, \hat{x}_{T|T-1}\} = L_T(p(x_0, x_1, \ldots x_T \mid y_0, y_1, \ldots, y_{T-1})).$$

[2] An equivalent assumption is that the disturbances vectors w_k are independent.

A typical choice for the functional $L_T(\cdot)$ is either an expectation or the maximum *a posteriori* Bayesian (MAP) estimate

$$\begin{aligned}&\{\hat{x}_{0|T-1}, \hat{x}_{1|T-1}, \dots, \hat{x}_{T|T-1}\} \in \\ &\quad \arg \max_{\{x_0, x_1, \dots, x_T\}} p(x_0, x_1, \dots, x_T \mid y_0, \dots, y_{T-1}).\end{aligned} \tag{2.5}$$

In this work we focus solely on the Bayesian criterion.

Solving (2.5) requires an expression for the conditional probability density function (2.4). Following the developments of Cox (1964) and Jazwinski (1970), we determine the conditional probability density function (2.4) as follows. We can express, using the Markov property, the joint probability of the state as

$$p(x_0, \dots, x_T) = p_{x_0}(x_0) \prod_{k=0}^{T-1} p(x_{k+1}|x_k),$$

where $p_{x_0}(x_0)$ denotes our prior information concerning the initial state of the system. If we assume the measurement noise v_k is independent, then using our model of the sensor (2.2) we have the relationship

$$p(y_0, \dots, y_{T-1} \mid x_0, \dots, x_{T-1}) = \prod_{k=0}^{T-1} p_{v_k}(y_k - h(x_k)).$$

Applying Bayes' rule, we obtain

$$p(x_0, x_1, \dots, x_T \mid y_0, \dots, y_{T-1}) \propto p_{x_0}(x_0) \prod_{k=0}^{T-1} p_{v_k}(y_k - h(x_k)) p(x_{k+1}|x_k).$$

The properties of logarithms allow us to establish following equality

$$\begin{aligned}&\arg \max_{\{x_0, x_1, \dots, x_T\}} p(x_0, x_1, \dots, x_T \mid y_0, \dots, y_{T-1}), \\ &\quad = \arg \max_{\{x_0, x_1, \dots, x_T\}} \log p(x_0, x_1, \dots, x_T \mid y_0, \dots, y_{T-1}), \\ &\quad = \arg \max_{\{x_0, x_1, \dots, x_T\}} \sum_{k=0}^{T-1} \log p_{v_k}(y_k - h(x_k)) + \log p(x_{k+1}|p_k) + \log p_{x_0}(x_0).\end{aligned}$$

The last equation is useful, because it allows us to transform the problem (2.5) into a multi-stage optimization. As we illustrate, compression is conceptually easier to address when the problem structure is multi-stage.

We have succeeded in transforming the state estimation problem into a multi-stage dynamic optimization, though the formulation is still requires the specification of the probability density functions. The probability density functions p_{v_k} and p_{x_0} are commonly chosen as normals. Even though the choice is justified typically by the law of large numbers, one chooses normals, more often than not, because they are mathematically convenient. Evaluating the state transition probability density function $p(x_{k+1}|x_k)$, however, requires the solution of functional difference equation, the discrete-time analogue of the Kolmogorov's forward equation, unless we make the following simplifying assumptions[3]:

A The disturbances w_k are mutually independent;

[3]If the vector w_k is a normally distributed random variable, then we can replace assumption **B** with

$$f_d(x_k, u_k, w_k, k) = f(x_k, u_k, k) + G w_k,$$

where G is a matrix with full column rank.

B $f_d(x_k, u_k, w_k, k) = f(x_k, u_k, k) + w_k$.

Under these two assumptions, we have

$$p(x_{k+1} \mid x_k) = p_{w_k}(x_{k+1} - f(x_k, u_k, k)).$$

The probability density function $p_{w_k}(\cdot)$ is also commonly chosen as a normal. Assumptions **A** and **B** allow us to cast (2.5) as an optimization explicitly in terms of the process model and the probability density functions $p_{v_k}(\cdot)$, $p_{w_k}(\cdot)$, and $p_{x_0}(\cdot)$:

$$\begin{aligned}\arg \max_{\{x_0, x_1, \dots, x_T\}} & p(x_0, x_1, \dots, x_T \mid y_0, \dots, y_{T-1}) = \\ & \arg \max_{\{x_0, x_1, \dots, x_T\}} \sum_{k=0}^{T-1} \log p_{v_k}(y_k - h(x_k)) + \log p_{w_k}(x_{k+1} - f(x_k, u_k, k)) \\ & + \log p_{x_0}(x_0).\end{aligned}$$

If we assume furthermore that the density $p_{x_0}(\cdot)$ is normal with mean $\hat{x}_0$ and covariance Π_0 and the densities $p_{w_k}(\cdot)$ and $p_{v_k}(\cdot)$ are normal with zero mean and covariances Q and R respectively, then we have

$$\begin{aligned}\arg \max_{\{x_0, x_1, \dots, x_T\}} & p(x_0, x_1, \dots, x_T \mid y_0, \dots, y_{T-1}) = \\ & \arg \min_{\{x_0, x_1, \dots, x_T\}} \sum_{k=0}^{T-1} \|y_k - h(x_k)\|^2_{R^{-1}} + \|x_{k+1} - f(x_k, u_k, k)\|^2_{Q^{-1}} \\ & + \|x_0 - \hat{x}_0\|^2_{\Pi_0^{-1}},\end{aligned}$$

where $\|z\|^2_A = z^T A z$.

The normality assumptions are sufficient for many problems. However, we can improve our descriptions of the random variables w_k, v_k, and x_k by introducing the constraints

$$w_k \in \mathbb{W}_k, \qquad v_k \in \mathbb{V}_k, \qquad x_k \in \mathbb{X}_k,$$

where the sets $\mathbb{W}_k$, $\mathbb{V}_k$, and $\mathbb{X}_k$ are closed and convex. One commonly chooses the sets as finite-dimensional polyhedral convex sets: i.e.

$$\mathbb{W}_k = \{w_k \; : \; w^k_{\min} \le W_k w_k \le w^k_{\max}\}.$$

In a probabilistic framework, the constraint sets provide the support for the probability density functions. If, for example, we assume

$$\mathbb{W}_k = \{w_k \; : \; -1 \le w_k \le 1\},$$

and the probability density function $p_{w_k}(\cdot)$ is a normal with zero-mean and unit variance, then the constraints project the probability density function $p_{w_k}(\cdot)$ onto $\mathbb{W}_k$ yielding a truncated normal (see Figure 2.1). One obtains similar results if the probability density function $p_{v_k}(\cdot)$ is coupled with constraints. However, we advise against constraining the vector v_k due to the possibility of outliers. Constraints may amplify the affect of spurious measurements; if one constrains the measurement residual v_k, then the estimate $\hat{x}_{k|T-1}$ may be unable to ignore the spurious measurement y_k. One may also use constraints to generate asymmetric distributions by piecing together truncated probability density functions as a jigsaw using variable decompositions (Robertson 1996, Robertson and Lee 1998).

The probabilistic interpretation and implication of constraints on the state x_k is not as simple. Some of the issues are illustrated in the following simple example. Suppose we have a leaky vessel

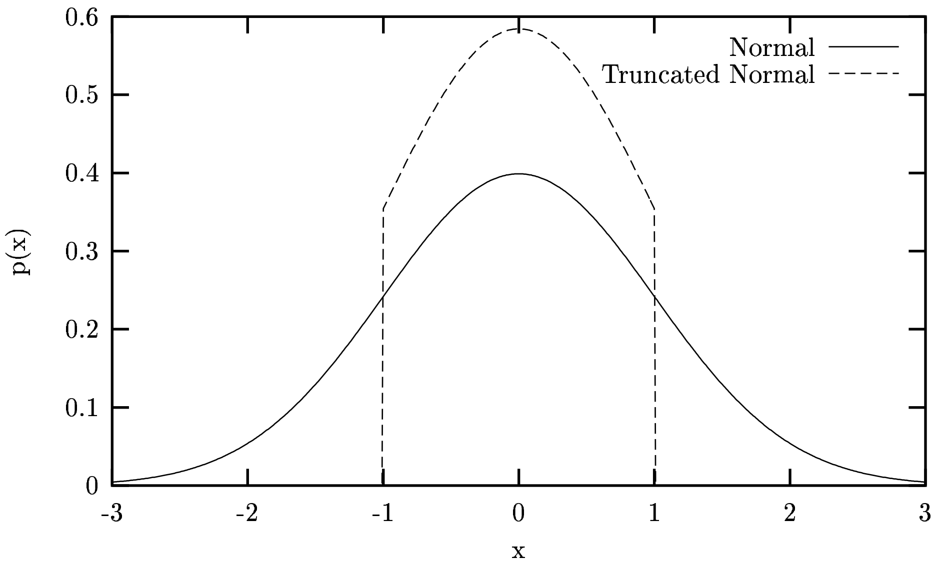


Figure 2.1: Comparison of a normal and truncated normal probability density function.

initially full of a liquid compound A. Let the state x_k denote the mass of A at time k and the vector w_k denote the mass of A that leaks from the vessel during the time interval k to $k+1$. A simple mass balance yields the model

$$x_{k+1} = x_k + w_k.$$

In addition to the mass balance, we know the state x_k is positive and bounded and the disturbance w_k is negative. One immediate consequence of the state constraint $x_k \geq 0$ is that the state x_k and disturbance w_k are correlated: if the state x_k is small, then the state constraint $x_k \geq 0$ implies the disturbance w_k is also necessarily small. This result is physically obvious, yet also somewhat surprising. One typically assumes that the exogenous disturbances are independent of the state of the process. If we ignore the effect of recycle and feedback loops, the disturbances are a result of variations in upstream processes unaffected by the state of the downstream process. Another consequence of state constraints is the violation of causality. If we rewrite the state equation explicitly in terms of the vector w_k, then we have the equivalent representation

$$x_{k+1} = x_0 + \sum_{k=0}^{k} w_k.$$

If we suppose that the initial leak w_0 is large, then the future leaks $\{w_1, w_2, \dots\}$ are necessarily small: there is less mass in the vessel that can leak out. Likewise, a large leak at time k requires that past leaks $\{w_0, w_1, \dots, w_{k-1}\}$ are small. This causal correlation is equivalent to the correlation between the disturbance vector w_k and the state x_k, because we model the system as a Markov process. Again, one commonly assumes the disturbances are mutually independent, and in this case they are not. The conclusions from this example are that state constraints may significantly alter the probabilistic structure of the problem. Rarely is this structure explicitly specified in the problem statement, so one should exercise care with state constraints. The advantage of state constraints is that they allow for simplified models: rather than having to develop a correlation between the mass in the vessel x_k and the leak w_k, we can use a simple mass balance in conjunction with constraints. Simplifying the modeling requirements is important, because the most time consuming task in design is model development (Ogunnaike 1995). We discuss the issue of constraints further in Sections 2.4 and 3.2.2.

Moving Horizon Estimation

Consider again the problem (2.5). Under the assumptions of normality, we can recast the state estimation problem at time T as the following mathematical program[4]:

$$\min_{x_0, \{w_k\}_{k=0}^{T-1}} \Phi_T(x_0, \{w_k\}), \tag{2.6}$$

subject to

$$x_{k+1} = f(x_k, u_k, k) + w_k, \tag{2.7a}$$

$$y_k = g(x_k, k) + v_k, \tag{2.7b}$$

$$w_k \in \mathbb{W}_k, \qquad v_k \in \mathbb{V}_k, \qquad x_k \in \mathbb{X}_k, \tag{2.7c}$$

where

$$\Phi_T(x_0, \{w_k\}) = \sum_{k=0}^{T-1} \|v_k\|^2_{R^{-1}} + \|w_k\|^2_{Q^{-1}} + \|x_0 - \hat{x}_0\|^2_{\Pi^{-1}}.$$

If we let $\hat{x}_{0|T-1}$ and $\{\hat{w}_{k|T-1}\}_{k=0}^{T-1}$ denote the solution to (2.6), then the optimal state estimate at time T is the sequences $\{\hat{x}_{k|T-1}\}_{k=0}^{T}$ obtained by solving the state equation (2.7a). The matrices Q and R, in this formulation, are the tuning parameters for reconciling the model with the process measurements. The matrices provide the means by which the errors are distributed between the model and the process data. In addition to their statistical significance, the matrices have the following simple interpretation: the matrix Q provides a measure of confidence in the model while the matrix R provides a measure of confidence in the process data. Thus, if the matrix Q is "large" relative to R, then we are less confident in the model than in the process data and vice-versa.

Many different options exist for solving the mathematical program (2.6)-(2.7). The problem as formulated requires the solution of a nonlinear program, a computationally demanding though tractable problem. If the process model is stiff or has unstable dynamics, a simultaneous strategy, in which the discretization and optimization are performed simultaneously, is often advantageous (Biegler 1997, Biegler 1998, Bock, Diehl, Leineweber and Schlöser 1998). When the process model is linear and the constraints are polyhedral convex sets, the mathematical program reduces to a quadratic program, a far less computationally demanding problem. Regardless of the problem complexity, solving the state estimation problem (2.5) online is usually impossible, because the size of problem (2.6)-(2.7) grows without bound as we collect more process measurements. Online implementation, therefore, requires that we bound the size of the mathematical program (2.6)-(2.7). Consequently we need a strategy to compress the data. The strategy we employ is *approximate* dynamic programming.

Consider the objective function $\Phi_T(\cdot)$. We may rearrange the objective function $\Phi_T(\cdot)$ by breaking the time interval into two pieces $t_1 = \{k : 0 \le k \le T-N-1\}$ and $t_2 = \{k : T-N \le k \le T-1\}$ as follows:

$$\begin{aligned}\Phi_T(x_0, \{w_k\}_{k=0}^{T-1}) \quad = \quad & \sum_{k=N-M}^{T-1} \|v_k\|^2_{R^{-1}} + \|w_k\|^2_{Q^{-1}} + \\ & \sum_{k=0}^{T-M-1} \|v_k\|^2_{R^{-1}} + \|w_k\|^2_{Q^{-1}} + \|x_0 - \hat{x}_0\|^2_{\Pi^{-1}}.\end{aligned}$$

[4] $\{w_k\}_{k=0}^{T-1} := \{w_0, w_1, \ldots, w_{T-1}\}$

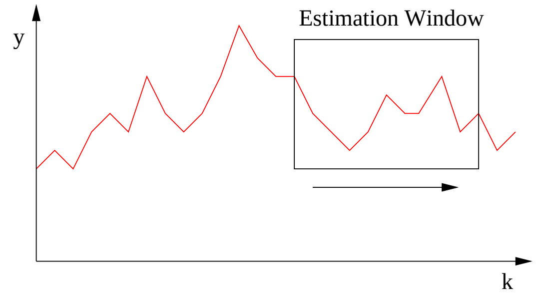


Figure 2.2: Graphical depiction of the moving horizon strategy.

Because we use a state-variable description of the system (i.e. a Markov process), the quantity

$$\sum_{k=N-M}^{T-1} \|v_k\|^2_{R^{-1}} + \|w_k\|^2_{Q^{-1}}$$

depends only on the state x_{T-N}, disturbance sequence $\{w_k\}_{k=T-N}^{T-1}$, and the process measurements $\{y_k\}_{k=T-N}^{T-1}$. The *principle of optimality* allows us to cast the estimation problem (2.5) as a fixed-horizon estimator. Standard dynamic programming arguments allow us to replace the mathematical program (2.6)-(2.7) with the following **equivalent** mathematical program

$$\min_{x_{N-M},\{w_k\}_{k=T-N}^{T-1}} \sum_{k=N-M}^{T-1} \|v_k\|^2_{R^{-1}} + \|w_k\|^2_{Q^{-1}} + \mathcal{Z}_{T-N}(x_{T-N})$$

subject to the constraints (2.7) where

$$\mathcal{Z}_\tau(p) = \min_{x_0,\{w_k\}_{k=0}^{\tau-1}} \{\Phi_\tau(x_0,\{w_k\}) \ : \ x_\tau = p\} \tag{2.8}$$

subject to the constraints (2.7). The mathematical program (2.8) provides the general structure for moving horizon estimation. Whereas in the problem (2.6)-(2.7) we considered all of the available process measurements, in moving horizon estimation we account explicitly only for the last N process measurements. We account for the remaining process measurements using the function $\mathcal{Z}_{T-N}(\cdot)$. The name "moving horizon estimation" arises from the analogy of a sliding estimation window (see Figure 2.2).

We refer to the function $\mathcal{Z}_\tau(\cdot)$ as the **arrival cost**. Arrival cost is fundamental in estimation, because, by providing a means to compress the data, it allows us to transform the unbounded mathematical problem into an equivalent fixed-dimension mathematical program. The arrival cost compactly summarizes the effect of the data $\{y_k\}_{k=0}^{\tau-1}$ on the state x_τ, thereby allowing us to fix the dimension of the optimization. We can view arrival cost as the analogue of the "cost to go" in standard backward dynamic programming. In probabilistic terms, the arrival cost generates the conditional density function $p(x_\tau|y_0,\dots,y_{\tau-1})$ and vice-versa: the arrival cost is proportional to the negative logarithm of the conditional density function $p(x_\tau|y_0,\dots,y_{\tau-1})$. Hence, we may view arrival cost as an equivalent statistic (Striebel 1965) for the conditional density function $p(x_\tau|y_0,\dots,y_{\tau-1})$. Further discussion on the properties of arrival cost may be found in Chapter 3.

Arrival cost provides a general method for compressing the process data. An excellent example of arrival cost is the Riccati equation arising in Kalman filtering. Consider the problem (2.6)-(2.7) where we assume the model is linear

$$x_{k+1} = Ax_k + w_k, \qquad y_k = Cx_k + v_k$$

and ignore the constraints $\mathbb{X}_k$, $\mathbb{W}_k$, and $\mathbb{V}_k$. If we use the Kalman filter covariance update formula

$$\Pi_T = GQG^T + A\Pi_{T-1}A^T - A\Pi_{T-1}C^T(R + C\Pi_{T-1}C^T)^{-1}C\Pi_{T-1}A^T \tag{2.9}$$

subject to the initial condition $\Pi_0 = \Pi$, then we can express the arrival cost explicitly as

$$\mathcal{Z}_T(x) = (x - \hat{x}_T)^T \Pi_T^{-1}(x - \hat{x}_T) + \Phi_T^*,$$

where $\hat{x}_T$ denotes the optimal estimate at time T given the measurements $\{y_k\}_{k=0}^{T-1}$ and Φ_T^* denotes the optimal cost at time T (see Appendix D for the derivation). From the preceding arguments, we have

$$\begin{aligned}
&\min_{x_0,\ \{w_k\}_{k=0}^{T-1}} \Phi_T(x_0,\ \{w_k\}) \equiv \\
&\min_{x_{T-N},\ \{w_k\}_{k=T-N}^{T-1}} \sum_{k=T-N}^{T-1} \|v_k\|^2_{R^{-1}} + \|w_k\|^2_{Q^{-1}} + \\
&\|x_{T-N} - \hat{x}_{T-N}\|^2_{\Pi^{-1}_{T-N}} + \Phi^*_{T-N}.
\end{aligned}$$

We can extract the Kalman filter by considering a horizon of $N = 1$. For this scenario, we have

$$\begin{aligned}
\Phi_T(x_{T-1}, w_{T-1}) = {}& v_{T-1}^T R^{-1} v_{T-1} + w_{T-1}^T Q^{-1} w_{T-1} + \\
& (x_{T-1} - \hat{x}_{T-1})^T \Pi_{T-1}^{-1}(x_{T-1} - \hat{x}_{T-1}).
\end{aligned}$$

Substituting in the model equations, evaluating the minimum with respect to w_{T-1} and x_{T-1}, and using some algebra, we obtain the well known result

$$\hat{x}_T = A\hat{x}_{T-1} + L(y_T - CA\hat{x}_{T-1})$$

for the Kalman filter, where

$$L = A\Pi_{T-1}C^T(R + C\Pi_{T-1}C^T)^{-1}.$$

Algebraic expressions for arrival cost do not exist unfortunately when either constraints are present or the process model is nonlinear. As these are the problems of interest, we need to generate *approximate* algebraic expressions for the arrival cost. At one extreme, we can discard the past information by approximating the arrival cost as a constant function: $\hat{\mathcal{Z}}_\tau(\cdot) = \Phi^*_\tau$. At the other extreme, we can ignore the current measurements and consider only the past measurements by approximating the arrival cost with the extended real-valued function

$$\hat{\mathcal{Z}}_\tau(x_\tau) = \begin{cases} \Phi^*_\tau & : \quad x_\tau = \hat{x}_\tau \\ \infty & : \quad x_\tau \neq \hat{x}_\tau \end{cases}.$$

Both of these choices are undesirable. Rarely are we completely ignorant or informed of the value of the state x_τ. One strategy to approximate the arrival cost is to use a first-order Taylor series approximation of the model around the estimated trajectory $\{\hat{x}_k\}_{k=0}^{\tau}$. This strategy approximates the arrival cost with an extended Kalman filter covariance update formula. We interpret this strategy as a neighboring extremal paths strategy in the context of estimation. Neighboring extremal paths are used to generate approximate optimal feedback laws for nonlinear systems by employing an extended linearization (Bryson and Ho 1975). The basic idea is as follows. If the deviation from the optimal path is small, then a linear approximation at the optimal path accurately describes the neighboring path.

If we let

$$A_k = \left.\frac{\partial f(x_k, u_k, k)}{\partial x_k}\right|_{\hat{x}_T}, \qquad C_k = \left.\frac{\partial g(x_k)}{\partial x_k}\right|_{\hat{x}_T},$$

then we obtain the extended Kalman filter covariance recursively from the equation

$$\Pi_{T+1} = Q + A_T(\Pi_T - \Pi_T C_T^T (R + C_T \Pi_T C_T^T)^{-1} C_T \Pi_T) A_T^T$$

subject to the initial condition $\Pi_0 = \Pi$. The choice

$$\hat{\mathcal{Z}}_\tau(x) = \|x - \hat{x}_\tau\|^2_{\Pi_\tau^{-1}} \tag{2.10}$$

summarizes our best available knowledge, to a first-order approximation, without introducing extra knowledge not available from the measurements. Using the extended Kalman filter to approximate the arrival cost has many advantages. When there are no constraints, one can view the estimator as an iterated extended Kalman filter. When the process model is linear, the estimator reduces to a Kalman filter.

One needs to wary of divergence (instability) when approximating the arrival cost. So long as the approximate arrival cost $\hat{\mathcal{Z}}_\tau(\cdot)$ satisfies certain technical conditions, one is guaranteed non-divergence, or stability (see Chapter 3). When the process model is linear, the Kalman filter covariance, regardless of whether there are constraints, yields a stable estimator (see Chapter 4). However, when the process model is nonlinear, the extended Kalman filter covariance does not guarantee stability, and additional measures are necessary to guarantee stability. In practical terms, there should be a degree of forgetting: the estimator should not weight the past data too heavily. One property of the Kalman filter is that it exponentially forgets the past data (c.f. Anderson (1999)). If one is concerned about estimator divergence, then adding a "forgetting factor" to the approximate arrival cost improves the estimator's "robustness". A simple strategy for generating a forgetting factor is to pre-multiply the approximate arrival cost by a scalar $\alpha \in (0, 1)$:

$$\hat{\mathcal{Z}}_T(x) = \alpha \|x - \hat{x}_T\|^2_{\Pi_T^{-1}}.$$

The interested reader is directed to Chapter 3 for further discussion regarding forgetting factors in constrained moving horizon estimation.

We formulate, therefore, moving horizon estimation (MHE) at time T as the solution to the following mathematical program

$$\min_{x_{T-N}, \{w_k\}_{k=T-N}^{T-1}} \hat{\phi}_T(x_{T-N}, \{w_k\}) \tag{2.11}$$

subject to the constraints

$$\begin{aligned} x_{k+1} &= f(x_k, u_k, k) + w_k, \\ y_k &= g(x_k, k) + v_k, \\ w_k \in \mathbb{W}_k, \qquad & v_k \in \mathbb{V}_k, \qquad x_k \in \mathbb{X}_k, \end{aligned}$$

where

$$\hat{\phi}_T(x_{T-N}, \{w_k\}) = \sum_{k=T-N}^{T-1} \|w_k\|^2_{Q^{-1}} + \|v_k\|^2_{R^{-1}} + \|x_{T-N} - \hat{x}_{T-N}\|^2_{\Pi_{T-N}^{-1}}.$$

If we let, with abuse of notation, $\hat{x}_{T-N|T-1}$ and $\{\hat{w}_{k|T-1}\}_{k=T-N}^{T-1}$ denote the solution to (2.11), then the optimal state estimate at time T is the sequences $\{\hat{x}_{k|T-1}\}_{k=T-N}^{T}$ obtained by solving the state equation (2.7a). Unlike the "full information" problem (2.6)-(2.7), the MHE estimator generates only truncated estimates–the consequence of considering only the data sequence $\{y_k\}_{k=T-N}^{T-1}$. For notation simplicity, let $\hat{x}_T := \hat{x}_{T|T-1}$

2.3.1 CSTR Example

To illustrate the ideas presented, consider a first order, exothermic, irreversible reaction $A \to B$ carried out in continuous stirred tank reactor (CSTR). For control and monitoring purposes, suppose we need to measure the time evolution of conversion. One solution is to measure directly the concentration of species A or B. While this solution is optimal, directly measuring the chemical species may be neither practical nor feasible due to operating conditions and economic factors. A more practical solution, at least in terms of the sensor requirements, is to infer the conversion from temperature. In particular, we seek an algorithm that reconstructs the conversion from the temperature measurements.

There are many ways to account for uncertainty in our model of the CSTR. For simplicity we assume there is only an additive, unmeasured, temperature disturbance ω_T. We may view the temperature disturbance ω_T as an aggregate disturbance encompassing such effects as feed temperature fluctuations, heat loss, and heat exchanger fouling. To account for uncertainty, we model the disturbance ω_T as a stochastic process. If we make standard simplifying assumptions such as perfect mixing, then we can model the CSTR as the following dimensionless stochastic differential equation

$$\frac{dx_1}{dt} = -x_1 + Da(1-x_1)\exp\left\{\frac{x_2}{1+x_2/\mathcal{E}}\right\}, \tag{2.12a}$$

$$\frac{dx_2}{dt} = -x_2 + BDa(1-x_1)\exp\left\{\frac{x_2}{1+x_2/\mathcal{E}}\right\} - U(x_2 - x_{2c}) + \omega_T, \tag{2.12b}$$

$$d\omega_T = a\omega_T dt + d\beta_t, \tag{2.12c}$$

where $\{\beta_t,\ t \geq 0\}$ is a Brownian motion process with $E\{d\beta_t^2\} = \delta Q dt$. The variable x_1 is the conversion, x_2 is the dimensionless temperature, t is the dimensionless time, and x_{2c} is the dimensionless cooling water temperature. The parameter B is the dimensionless adiabatic temperature rise, Da is the Damköhler number, $\mathcal{E}$ is the dimensionless activation energy, and U is the dimensionless heat transfer coefficient. See the article by Uppal, Ray and Poore (1974) for the details of the dimensionless variables. It is common to assume the measurements are corrupted with white noise. As reactor temperature is measured only at discrete times, we use the following model

$$y(t_k) = x_2(t_k) + v_k$$

to relate the measurements to the reactor temperatures, where the sequence $\{t_k\}_{k=0}^{\infty}$ denotes the times when measurements are available and the white noise sequence $\{v_k\}_{k=0}^{\infty}$ is used to represent measurement error. We assume for simplicity that $t_k = k\Delta t$, where Δt is the sampling period. The optimal Bayesian estimate for the conversion $\hat{x}_1(t_k)$ at time t_k is the most probable conversion given the sequence of temperature measurements $\{x_2(t_j)\}_{j=0}^{k}$. We formulate the problem as the solution to following optimization

$$\hat{x}_1(t_k) \in \arg\max_{x_1(t_k)} p(x_1(t_k) \mid x_2(t_0), \ldots x_2(t_k)).$$

To determine the conditional probability density function, we require expressions for the state transition probability density function $p(x_j|x_{j-1})$ and the measurement probability density function $p_{v_k}(\cdot)$. We can obtain an expression for the state transition probability density function, in theory, by solving Kolmogorov's forward equation. If we make the definition

$$f(x) := \begin{bmatrix} -x_1 + Da(1-x_1)\exp\left\{\dfrac{x_2}{1+x_2/\mathcal{E}}\right\} \\ -x_2 + BDa(1-x_1)\exp\left\{\dfrac{x_2}{1+x_2/\mathcal{E}}\right\} - U(x_2 - x_{2c}) + \omega_T \\ a\omega_T \end{bmatrix}$$

with the variable substitution

$$x := \begin{bmatrix} x_1 & x_2 & \omega_t \end{bmatrix}^T,$$

then Kolmogorov's forward equation becomes a partial differential equation of the form

$$\frac{\partial}{\partial t} p(x,t;y) = -\frac{\partial}{\partial x} \cdot [p(x,t;y)f(x)] + \frac{\partial^2}{\partial \omega_T^2} [p(x,t;y)\delta Q] \tag{2.13}$$

with boundary conditions

$$\lim_{t \to 0^+} p(x,t;y) = \delta(x-y)$$

and

$$p(\infty,t;y) = p(-\infty,t;y) = 0.$$

For a constant sampling period Δt, we obtain the equality

$$p(x_j|x_{j-1}) = p(x_j, \Delta t; x_{j-1}).$$

The cost of solving (2.13) numerically is far too large considering the arbitrary choice of the disturbance structure. So instead, we seek an approximate representation of the transition probability density function by linearizing Kolmogorov's forward equation. If we let

$$A := \frac{\partial}{\partial x} f(\hat{x})$$

for some $\hat{x}$, preferably a steady state, then we can approximate the state transition probability density function with the solution of the following continuous time Riccati equation

$$\frac{dQ(t)}{dt} = AQ(t) + Q(t)A^T + \delta Q e_3 e_3^T$$

subject to the initial condition

$$Q(0) = 0,$$

where e_3 is the unit vector whose 3rd entry is 1. The Riccati equation is obtained by taking the Fourier transform of (2.13) and recognizing the solution to (2.13) is the characteristic function of a normal distribution (Jazwinski 1970). Recall, normal distributions are invariant under linear transformation. Hence, we obtain the following approximate expression

$$p(x_j|x_{j-1}) \sim N(f_d(x_{j-1}), Q(\Delta t))$$

for the state transition probability distribution with a constant sampling period Δt. The function $f_d(x_j) := z(\Delta t, x_j)$, where $z(\cdot, x_j) : [0, \Delta t] \to \mathbb{R}^3$ is a continuous function that satisfies the integral equation

$$z(\tau, x_j) = x_j + \int_0^\tau f(z(\xi, x_j))\, d\xi.$$

If we assume that the measurement errors for the CSTR are zero mean and normally distributed with covariance σ_v^2, then we obtain the following expression for the measurement probability distribution

$$p_{v_k}(y_k - x_2(t_k)) \sim N(0, \sigma_v^2).$$

B	14.94	U	2.09
Da	1.00	δQ	5×10^{-5}
$\mathcal{E}$	25	a	-0.1
Δt	0.5	σ_v	1

Table 2.1: Dimensionless parameters for CSTR example

For simplicity and consistency, we assume

$$p(x_0) \sim N(\hat{x}_0, \Pi),$$

where $x(t_0)$ denotes our prior knowledge of the state. If we take logarithms and discretize the problem with a zero order hold on the Brownian motion increments $d\beta$, then the associated optimization becomes

$$\min_{x_0, \{w_k\}_{k=0}^{T-1}} \sum_{k=0}^{T-1} w_k^T Q^{-1}(\Delta t) w_k + \frac{\{y_k - x_k)\}^2}{\sigma_v^2} + (x_0 - \hat{x}_0)^T \Pi^{-1} (x_0 - \hat{x}_0)$$

where the systems dynamics are specified by the nonlinear difference equation

$$x_{k+1} = f_d(x_k) + w_k.$$

Simulation Results

We designed the moving horizon estimator using a the horizon length $N = 5$. The stochastic differential equation (2.12) was solved using a first order Euler technique (Kloeden and Platen 1992). The estimator used the optimization algorithm NPSOL (Gill, Murray, Saunders and Wright 1986) and differential equation solver LSODE. The parameters are listed in Table 2.1. Figure 2.3 shows a comparison of the true arrival cost and the extended Kalman filter approximation as a function of the conversion x_1 and the dimensionless temperature x_2. As the figure demonstrates, the extended Kalman filter update is an effective approximation for the arrival cost. Figures 2.4 and 2.5 shows a plot of the MHE for a particular realization of the stochastic process. Figures 2.4 and 2.7 shows a plot of the MHE for another realization of the stochastic process where an ignition in the CSTR occurs. As both figures demonstrate, the MHE does an effective job of reconstructing the state. Due to the relative simplicity of this example, the extended Kalman filter yields comparable performance.

2.4 Constraints

The *raison d'être* for moving horizon estimation (MHE) is the ability to incorporate constraints in estimation. One might plausibly argue, however, that nonlinear dynamics also motivate the use of MHE. Unlike many other estimation strategies, MHE provides stability guarantees (see Chapter 3). One may also construct a stable estimator using a local coordinate transformation by output injection (Bestle and Zeitz 1983, Krener and Isidori 1983). However, unlike differential geometric methods, moving horizon strategies are applicable to a large class of problems. In particular, any feedback linearizable system can be stabilized also with a moving horizon controller (Meadows, Henson, Eaton and Rawlings 1995). We expect a similar result holds for estimation. Stability guarantees are important, but performance is the predominant concern. The extended Kalman filter provides only weak local stability guarantees (c.f. (Song and Grizzle 1995)), yet performs as well as most other estimation strategies. A "folk" theorem in

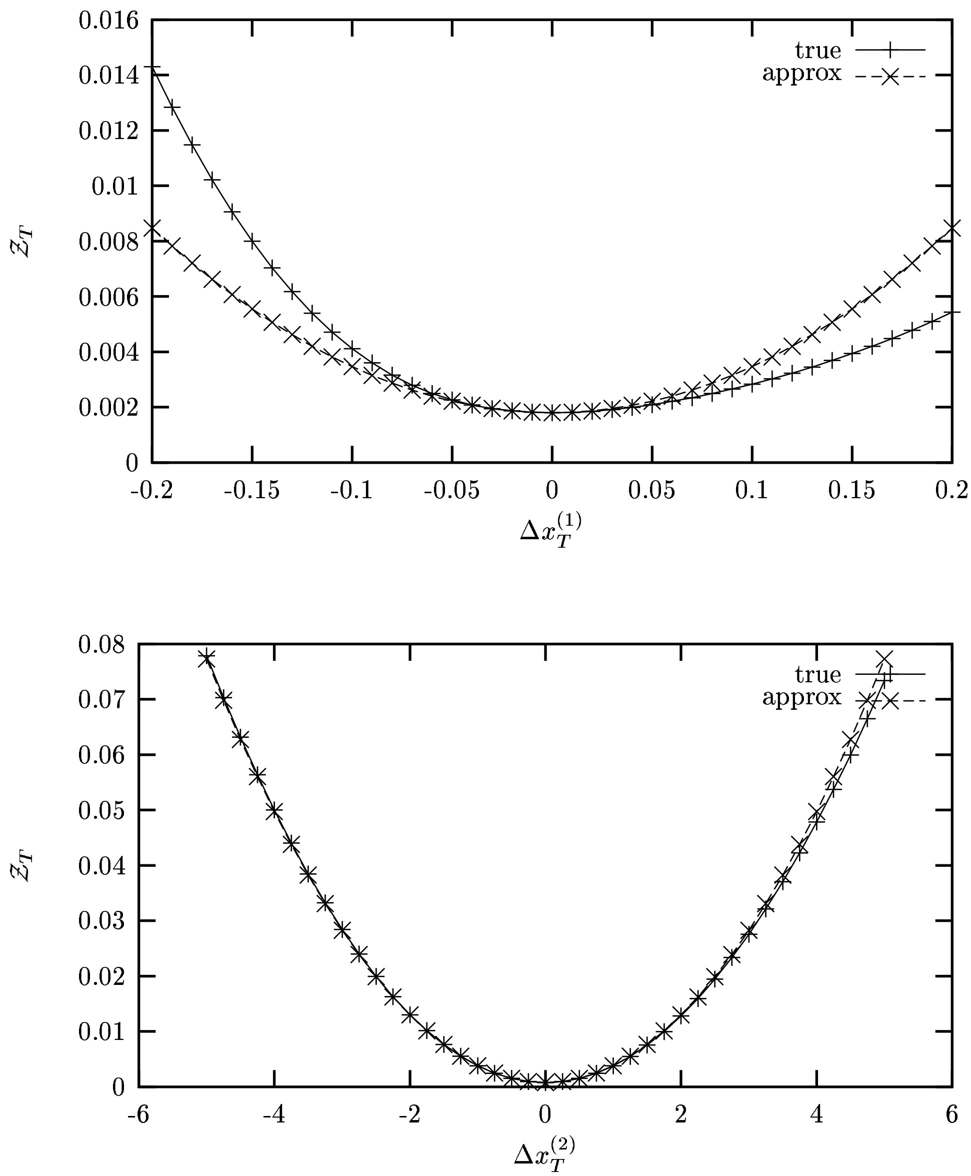


Figure 2.3: A comparison of the true arrival cost and extended Kalman filter approximation for CSTR example with $M = 5$.

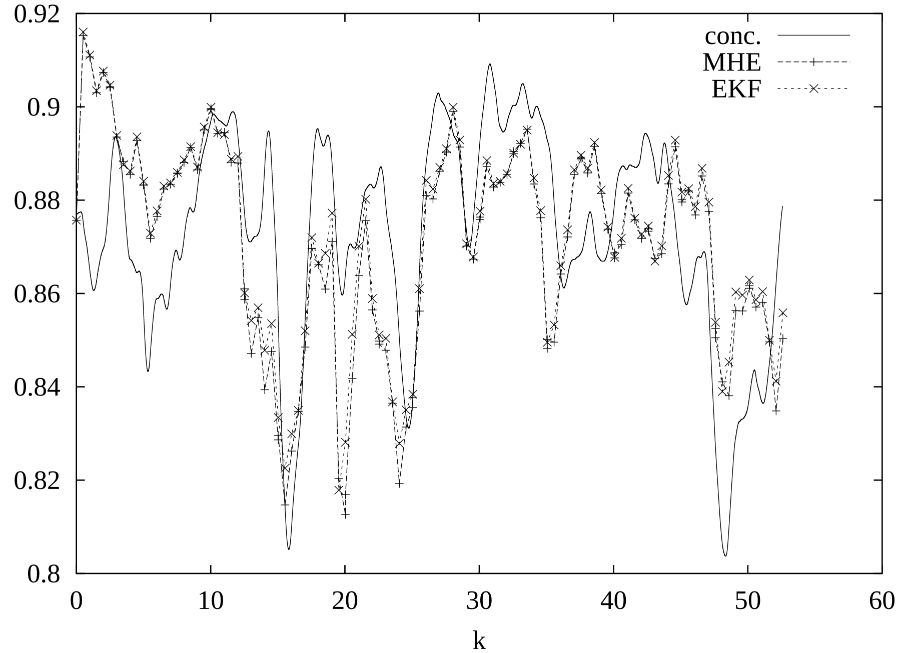


Figure 2.4: CSTR Example 1: Comparison of true and estimated conversion

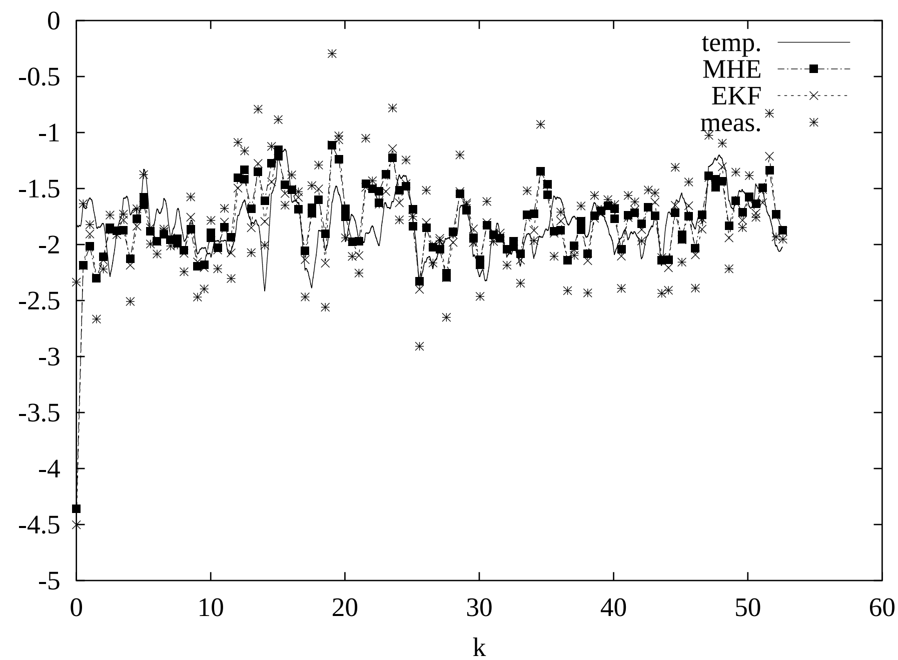


Figure 2.5: CSTR Example 1: Comparison of true and estimated temperature

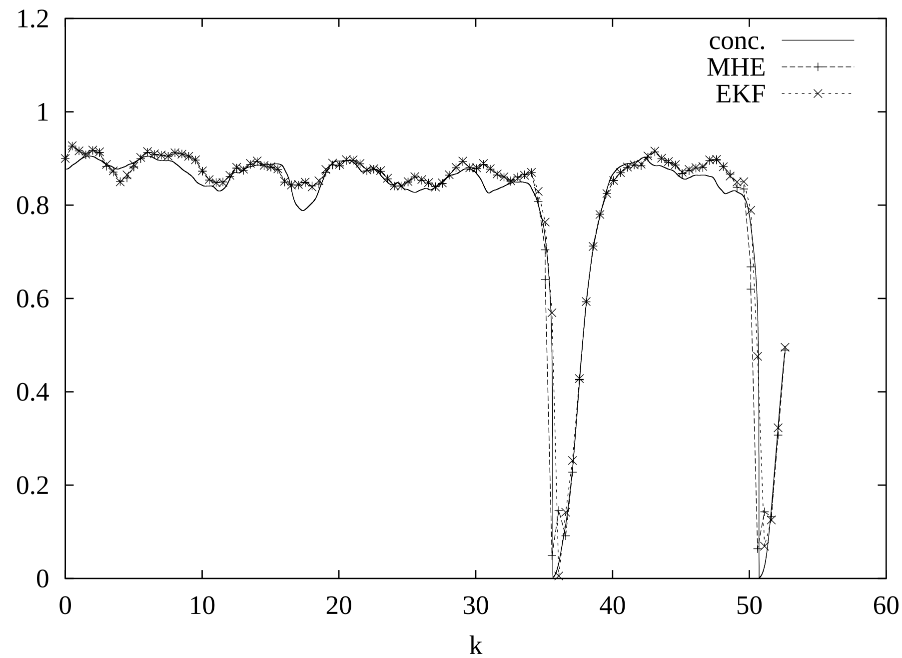


Figure 2.6: CSTR Example 2: Comparison of true and estimated conversion

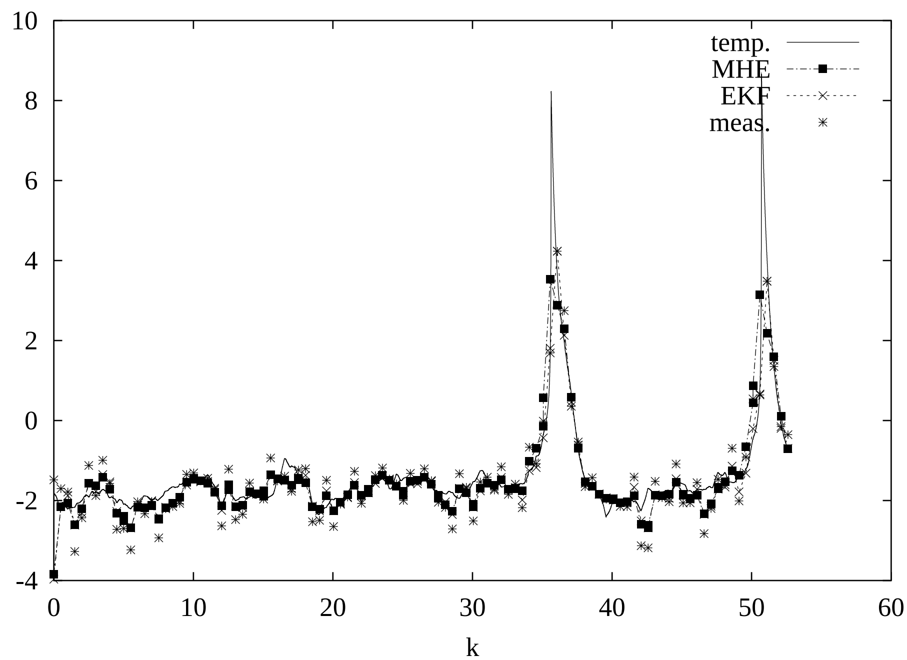


Figure 2.7: CSTR Example 2: Comparison of true and estimated temperature

control theory claims that if the noise is relatively small, then the performance of the extended Kalman filter is nearly optimal. Whereas, if the noise is large, then designing any monitoring or control strategy is "misguided" (Wonham 1969). The last thirty years have failed to convince us otherwise.

We may view unconstrained MHE as a form of extended Kalman filtering or, rather, the extended Kalman filter as a form of unconstrained MHE. The difference between the two strategies is the degree of optimization: the extended Kalman filter takes only one Newton step, while unconstrained MHE takes as many Newton steps as necessary to satisfy the (local) optimality conditions. We view, therefore, MHE as a form of iterated extended Kalman filtering and the extended Kalman filter as a suboptimal strategy for unconstrained MHE with a horizon length $N = 1$. One reason for the success of the extended Kalman filter is that often most of the cost reduction in optimization is obtained during the first few Newton steps. Performances rarely improves tangibly if one iterates further.

Constraints motivate, therefore, the use of MHE. We can best illustrate the potential of MHE with the following examples.

2.4.1 Example of inequality constraints yielding improved estimates.

Consider the following discrete-time system[5]

$$x_{k+1} = \begin{bmatrix} 0.9962 & 0.1949 \\ -0.1949 & 0.3815 \end{bmatrix} x_k + \begin{bmatrix} 0.03393 \\ 0.1949 \end{bmatrix} w_k, \qquad y_k = \begin{bmatrix} 1 & -3 \end{bmatrix} x_k + v_k. \tag{2.14}$$

We assume $\{v_k\}$ is sequence of independent, zero mean, normally distributed random variables with covariance 0.01, and $w_k = |z_k|$ where $\{z_k\}$ is a sequence of independent, zero mean, normally distributed random variables with unit covariance. We also assume the initial state x_0 is normally distributed with zero mean and covariance equal to the identity.

We formulate the constrained estimation problem with $Q = 1$, $R = 0.01$, $\Pi_0 = 1$, and $\hat{x}_0 = 0$. For the MHE, we choose $N = 10$. To capture our knowledge of the random sequence w_k, we add the inequality constraint $w_k \geq 0$. Note, this formulation yields the *optimal* Bayesian estimate. A comparison of the Kalman filter, full information estimator, and MHE for a single realization of (2.14) is shown in Figure 2.8. As expected, the performance of the constrained estimators is superior to the Kalman filter, because the constrained estimators possess, with the addition of the inequality constraints, the proper statistics of the disturbance sequence w_k. Hence, the constrained estimation problem formulated above accurately models the random variable w_k.

If we consider the statistics of the random variable w_k, it is important to note that the mean is not zero and the covariance is not 1. Rather, the mean is $2/\sqrt{2\pi}$ and the covariance is $(1 - 2/\pi)$. When we consider the negative inverse logarithm of the probability density function, however, we have

$$-\log p_{w_k}(w_k) \propto \frac{1}{2} w_k' w_k \qquad \text{for } w_k \geq 0.$$

Note, therefore, that constraints allow for non-Gaussian disturbances.

2.4.2 Leak detection and inventory estimation.

Consider the problem of detecting the location and magnitude of a leak in the waste water treatment process shown in Figure 2.9. We suppose the process is described by the following linear state space

[5]This state space system is a realization of the following system $G(s) = \frac{-3s+1}{s^2+3s+1}$ sampled with a zero-order hold and sampling time of 0.3.

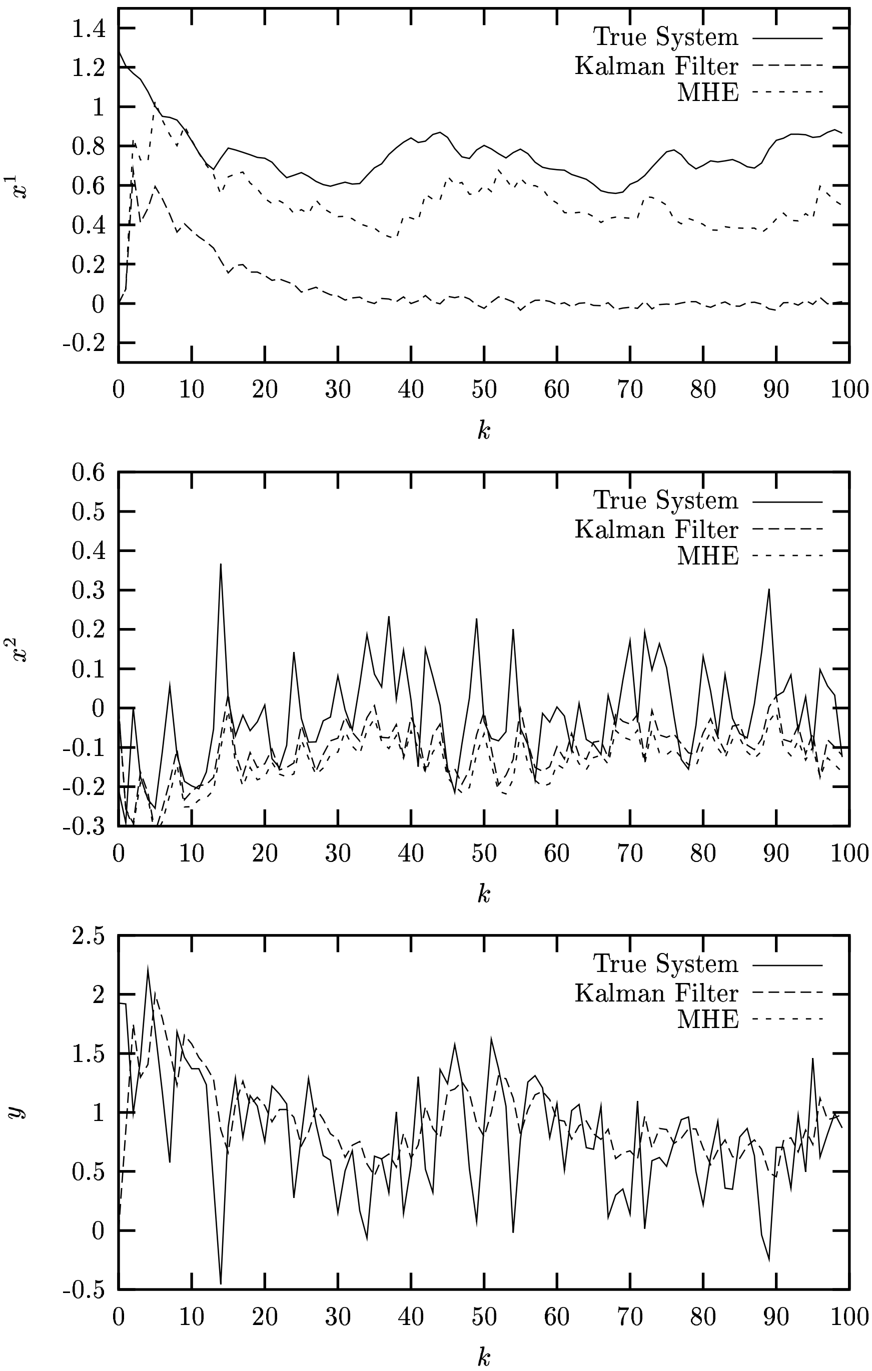


Figure 2.8: Comparison of estimators for Example 2.4.1.

]

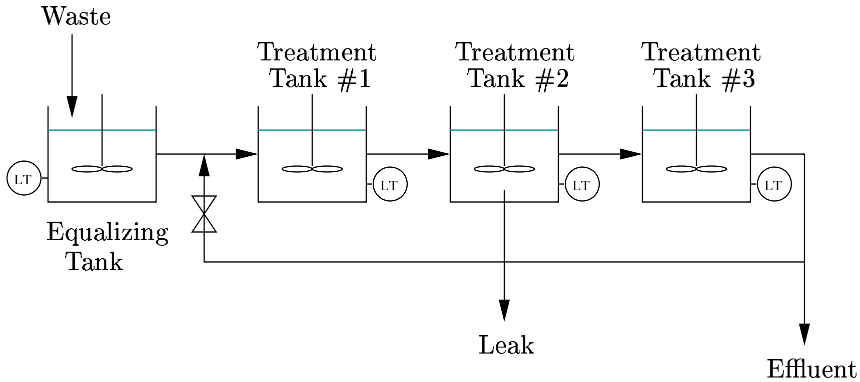


Figure 2.9: Schematic of tank process

model

$$x_{k+1} = \begin{bmatrix} 0.89168 & 0 & 0 & 0 & 1.0 \\ 0.10832 & 0.90518 & 0 & 0.04306 & 0 \\ 0 & 0.09482 & 0.89524 & 0 & 0 \\ 0 & 0 & 0.10476 & 0.89235 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} x_k + \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} w_k,$$

$$y_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & m \end{bmatrix} x_k + v_k.$$

The physical meaning of the entries in state variable x_k are given in Table 2.4.2. We choose $m = 1$ when we suppose the mass of waste entering the process is measured and $m = 0$ otherwise. We suppose the mass of each tank and the mass flowrate of waste entering the process are measured with the error covariance

$$R = \text{diag}\begin{bmatrix} 8 & 8 & 8 & 8 & 4 \end{bmatrix}.$$

As the leak is limited to waste tank #2, the process was simulated with $w_k = |z_k|$, where z_k is a normally distributed random variable with covariance matrix

$$Q_z = \text{diag}\begin{bmatrix} 0 & 0 & 5 & 0 & 15 \end{bmatrix}.$$

As the location of the leak is unknown (to the estimator), we design the estimator with the covariance matrix

$$Q = \text{diag}\begin{bmatrix} 5 & 5 & 5 & 5 & 15 \end{bmatrix}.$$

$x^{(1)}$	Mass in Equalizing Tank
$x^{(2)}$	Mass in Tank #1
$x^{(3)}$	Mass in Tank #2
$x^{(4)}$	Mass in Tank #3
$x^{(5)}$	Mass of Waste Entering Equalizing Tank

Table 2.2: State description for Example 2.4.2

Scenario		Total Losses	Mean Losses (by Tank)			
			Equal.	#1	#2	#3
Flow Measured	Actual	948.08	0	0	1.8962	0
	MHE	992.82	0.2429	0.2387	1.1253	0.3176
	KF	593.77	-0.2538	-0.0610	1.0262	0.0425
(No Leak)	Actual	0	0	0	0	0
	MHE	295.49	0.2418	0.2698	0.2925	0.27001
	KF	-186.44	-0.2552	-0.0129	0.0113	0.0014
Flow Unmeasured	Actual	916.81	0	0	1.8336	0
	MHE	907.54	0.1074	0.2427	1.0664	0.3123
	KF	405.24	-0.5722	-0.0626	0.9761	0.03860
(No Leak)	Actual	0	0	0	0	0
	MHE	244.87	0.1655	0.2729	0.2794	0.2648
	KF	-335.74	-0.5732	-0.0169	-0.0032	-0.0010

Table 2.3: Simulation results of Example 2.4.2

We, furthermore, added the constraints $w_k \geq 0$ and $x_k \geq 0$ in order to satisfy the mass balances: mass is only lost through a leak and the tanks must have positive mass. A horizon of $M = 10$ was chosen.

Two separate scenarios were considered (flow measured and unmeasured) along with a control where there is no leak. The results of the simulations are shown in Table 2.3. As one would expect, both the Kalman filter and MHE are able to detect the leak. The ability to detect the leak degrades when the flowrate is unmeasured. This result is expected as less information is available to both estimators. The benefit of constraints arise when one attempts to estimate the total losses. While MHE is able to provide a fairly accurate estimate of the total losses, the Kalman filter underestimates the total losses. The Kalman filter provides also *negative* estimates for the losses in the equalizing tank and tank # 1 in all four scenarios. Furthermore, when there is no leak, the Kalman filter predicts a net addition of mass to the tank system, which is, obviously, physically impossible. One can attribute this difference to the addition of constraints; the only difference between the two algorithms.

Note also that the constrained estimates are slightly biased away from zero in the tight tanks. Recall from the previous example that a truncated normal does not have a zero mean. Rather, the mean is $2\sigma/\sqrt{2\pi}$, where σ is the standard deviation of the corresponding normally distributed random variable. Any automated procedure involving hypothesis testing needs to account for this fact. This point is clearly illustrated when we simulate the waste treatment process without any leaks. When the flowrate is measured, the constrained estimate is worse than the unconstrained estimate. The reason for the poor estimates is due to the positive mean values: the constrained estimates have a mean bias of roughly 0.25. If we remove the bias, the estimate for the total leak is roughly zero as desired. However, if we remove the bias from the simulation where there is a leak in Tank #2, then the estimate for the total leak is the same as the Kalman filter. This example illustrates some of the issues one need be wary of when implementing constraints. While the constrained estimators provide a good estimate of the total losses when there is a leak, MHE and the Kalman filter both provide poor estimates when there are no leaks. The problem stems from an incorrect model of the process: the true process has no leaks, while the model assumes a leak in each tank. Nevertheless, one would normally use such a model in fault detection. Hence, any analysis would need to account for this discrepancy.

The "proper" strategy is to formulate this problem as a constrained signal detection problem. One would model all leak possibilities and then discriminate between the various scenarios using hypothesis testing. An alternative is to employ mixed integer programming (c.f. (Gatzke and Doyle III 1999)). As the focus of this chapter is not fault detection, but rather constrained monitoring, we do not pursue this topic further.

2.4.3 Semibatch Reactor

Consider the stirred-tank reactor depicted in Figure 2.14 where the following liquid phase exothermic reaction occurs

$$\mathrm{A} + 2\mathrm{B} \rightarrow \mathrm{C}.$$

The state estimation problem, inspired by the problem described by Rawlings, Jerome, Hamer and Bruemmer (1989), is to estimate precisely the concentration of A in the reactor. Because over addition of B leads to product degradation, precise concentration estimates of A as a function of time are necessary to complete the reaction without over addition of B. We suppose only temperature measurements corrupted with sensor noise are available. Furthermore, we suppose the exact reaction kinetics are unknown with the exception of the heat of reaction ΔH_r. The extent of reaction is estimated using reaction calorimetry (c.f. Schuler and Schmidt (1992)).

Under standard assumptions, such as negligible potential and kinetic energy effects, constant density, uniformly homogeneous mixture, and no phase transition, we simulated the reactor using the following model

$$\begin{aligned}
\dot{V} &= F, \\
\dot{A} &= -k_0 \exp\left(-\frac{ER}{T}\right) A\, B^2 - \frac{F}{V} A, \\
\dot{B} &= -2k_0 \exp\left(-\frac{ER}{T}\right) A\, B^2 + \frac{F}{V}(C_{Bf} - B), \\
\dot{T} &= -\frac{\Delta H_r}{\rho\, C_p} k_0 \exp\left(-\frac{ER}{T}\right) A\, B^2 + \frac{F}{V}(T_f - T) + \\
&\quad \frac{UA}{\rho\, C_p V}(T_c - T).
\end{aligned}$$

The model parameters are listed in Table 2.4. The flowrate profile, though scaled differently, is the one used in the operation of the industrial reactor described by Rawlings et al. (1989). To account for imperfect cooling and modeling inaccuracies, we assumed the cooling water temperature fluctuates. The flowrate profile and the cooling water temperature used in the simulation are shown in Figure 2.15. We

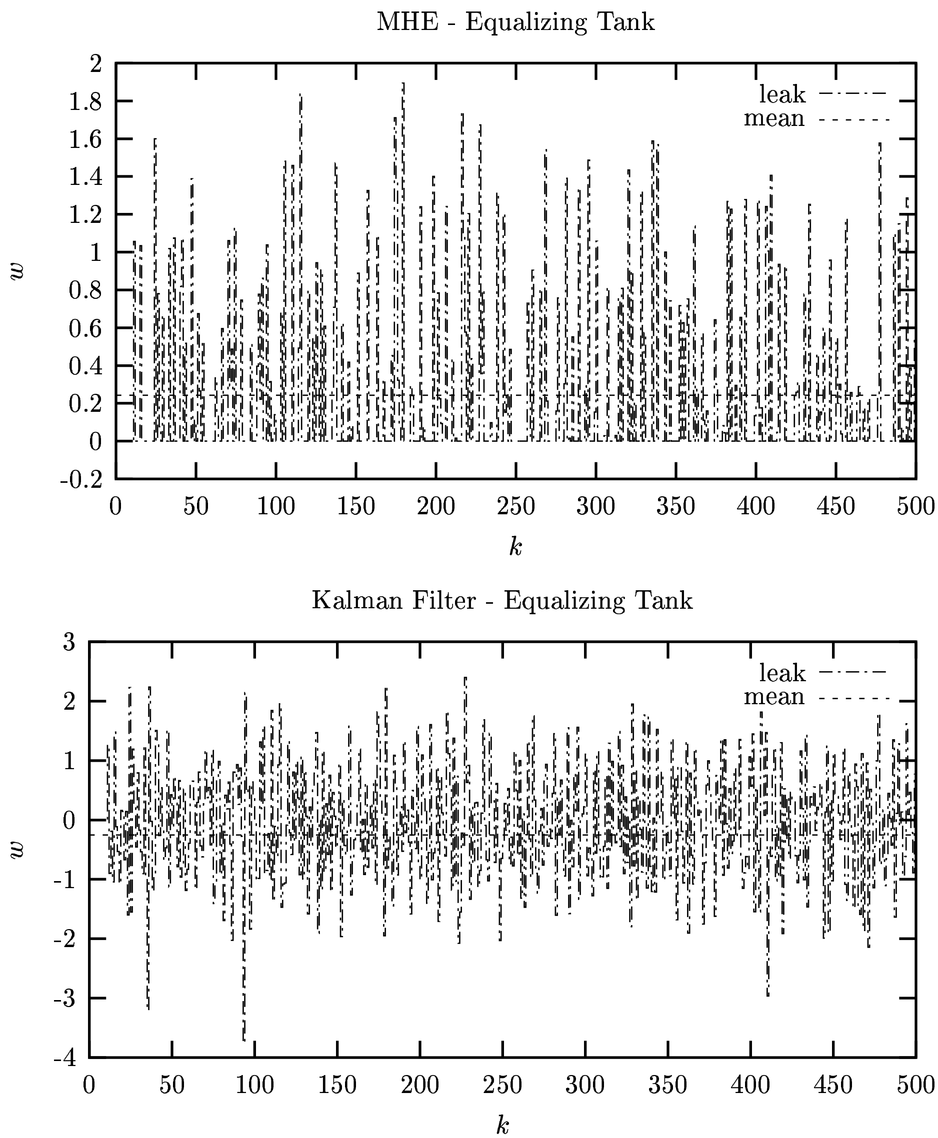


Figure 2.10: Leak estimates in the equalizing tank.

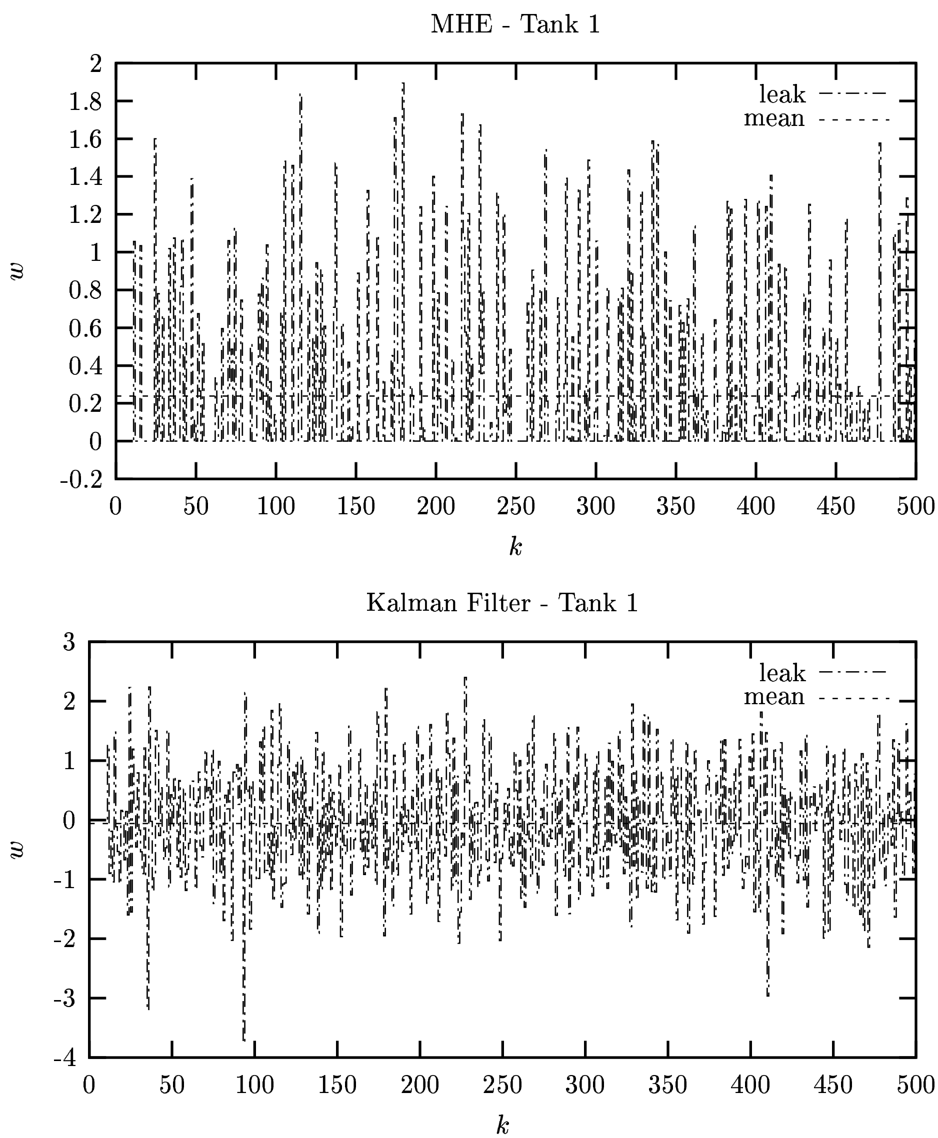


Figure 2.11: Leak estimates in the tank # 1.

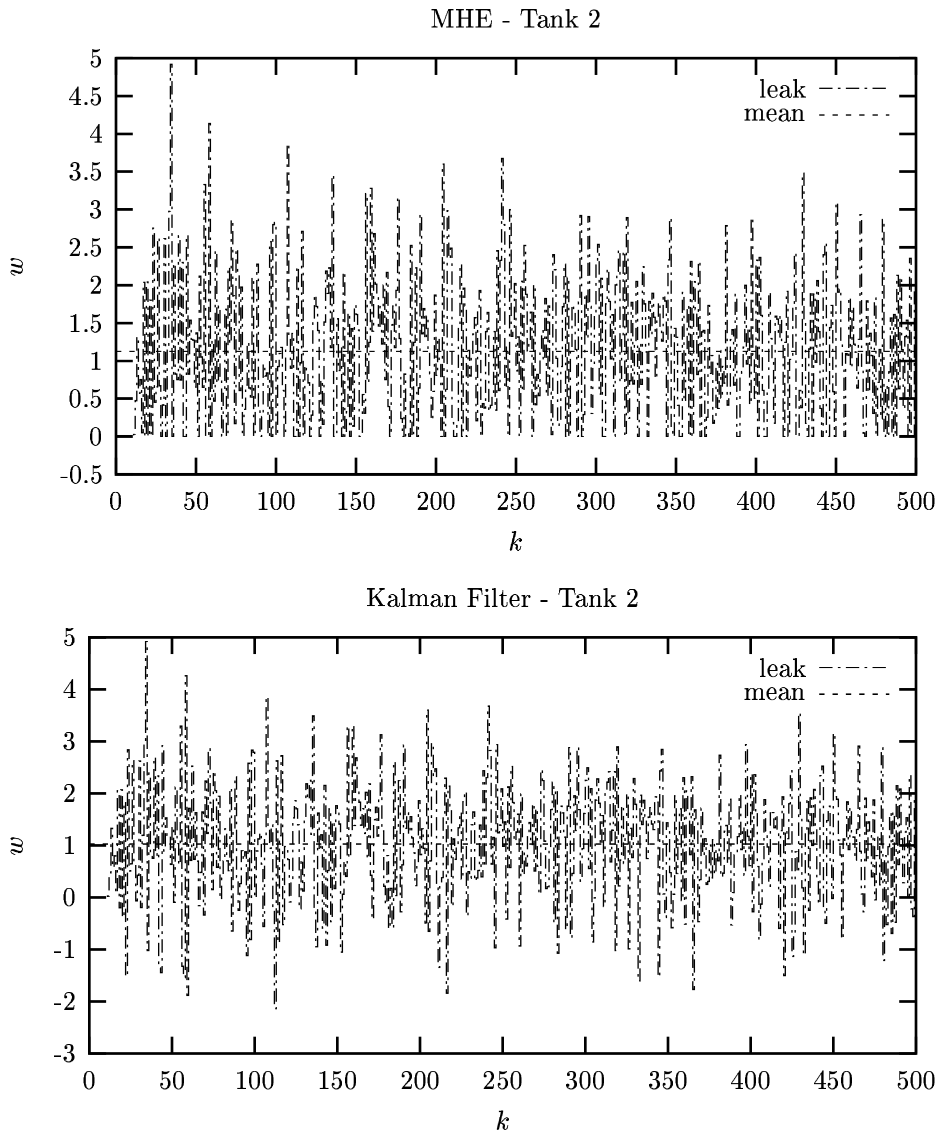
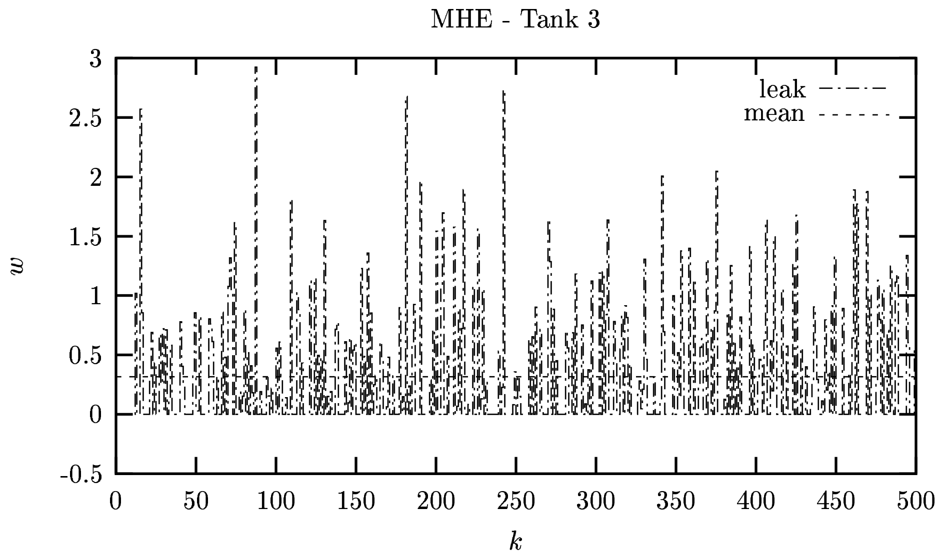


Figure 2.12: Leak estimates in the tank # 2.



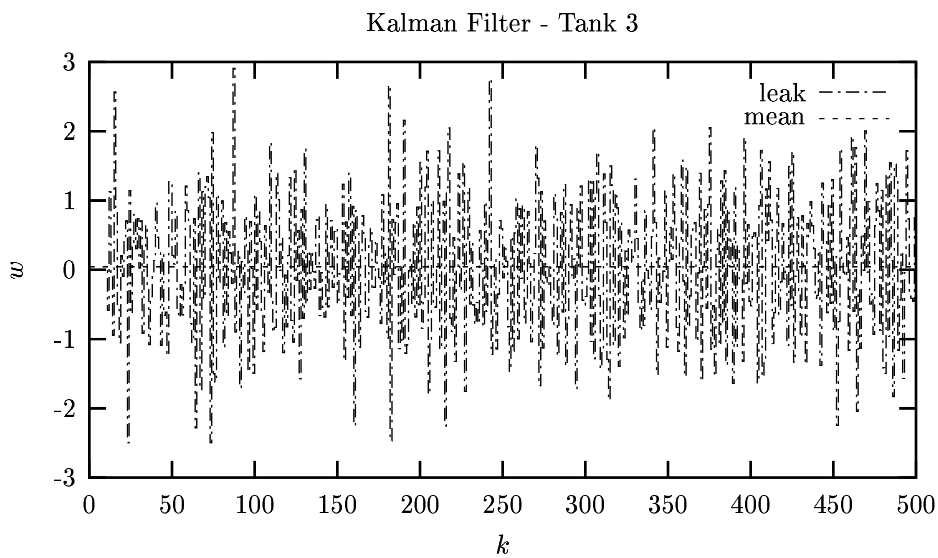


Figure 2.13: Leak estimates in the tank # 3.

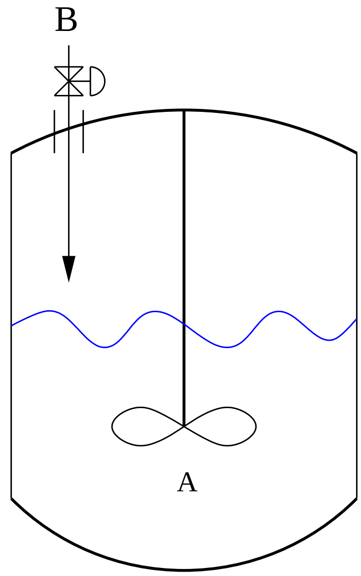


Figure 2.14: Reactor Schematic

k_0	$9e11$ mol^{-2} min^{-1}	$V(0)$	100 L
ER	$6e3$ K	$A(0)$	0.5 mol/L
ρ	1000 g/L	$B(0)$	0 mol
C_p	0.239 J/g·K	$T(0)$	300 K
UA	$2e5$ J/min·K		
T_f	300 K		
T_c	300 K		
C_{Bf}	2.2 mol/L		
$-\Delta H_r$	$5e4$ J/mol		

Table 2.4: Parameters for Example 2.4.3

suppose the temperature measurements are available every 30 seconds corrupted with zero mean and unit variance Gaussian noise.

The estimator has available only the following simplified time-varying linear model based on reaction calorimetry [6]:

$$\begin{aligned}
\dot{V} &= F, \\
\dot{A} &= r - \frac{F}{V}A, \\
\dot{B}(t) &= 2r + \frac{F}{V}(C_{Bf} - B), \\
\dot{T} &= \frac{\Delta H_r}{\rho\, C_p} r + \frac{F}{V}(T_f - T) + \frac{UA}{\rho\, C_p V}(T_c - T), \\
dQ_r &= dw.
\end{aligned}$$

[6] We also considered a model where the cooling water temperature fluctuations were included as a second disturbance. Our simulation results were no different.

The trick in reaction calorimetry is to estimate the reaction rate $r(\cdot)$ from the energy balance. The model was discretized with a zero-order hold and a sampling period of 30 seconds. The horizon length was $M = 10$

The advantage of the simplified model is that the reaction kinetics need not be known. However, as pointed out by (DeVallière and Bonvin 1990) and (M'hamdi, Helbig, Abel and Marquardt 1996), spurious estimates may result due to negative estimates of the reaction rate. We therefore constrain both the reaction rate and concentrations to be positive. We tuned the estimator with $Q = I$ and $\Pi_0 = I$ and initialized the estimator with the "true" initial conditions.

The results of the simulation are shown in Figure 2.16. Both the Kalman filter and MHE overestimate the actual reaction rate. This mismatch is due to fluctuations in the cooling water temperature. The addition of the constraints prevents MHE from estimating negative reaction rates and negative concentrations of A. Because MHE does not estimate negative reaction rates, the MHE estimate of reaction rate is larger than the Kalman filter estimate. Consequently, without the constraint on the concentrations, MHE would estimate also negative concentrations of A. The reason that the estimates are positive, even though the estimate of the reaction rate is too large, is due to smoothing. At each sampling time, MHE semi-implicitly estimates the entire reaction rate and concentration profile. We refer to these estimates as the smoothed estimates (i.e. $\hat{x}_{k|T}$ for $k \leq T$). The results shown in Figure 2.16 are only the tail of the estimated trajectory (i.e. $\hat{x}_T := \hat{x}_{T|T-1}$) and need not mutually satisfy the energy and mass balances. The smoothed estimates, however, mutually satisfy the energy and mass balances.

2.5 Conclusions

We have discussed moving horizon estimation (MHE) in the context of constrained process monitoring. MHE, as we have demonstrated through examples, is a practical and powerful strategy for constrained process monitoring. MHE allows the use of additional physical knowledge about systems, such as constraints and nonlinear dynamics, unavailable with other methods. While the ability to incorporate nonlinear dynamics is important, the distinguishing feature of MHE is the ability to incorporate inequality constraints. One can show, in particular, that MHE reduces to a Kalman filter or iterated extended Kalman filter when constraints are not present. Hence, we may view MHE as an extension of Kalman filtering.

Inequality constraints arise in many different contexts. We have illustrated the importance of inequality constraints in the following situations.

Truncated Distributions One often possesses prior knowledge in the form of bounds on the disturbances, state variables, and unknown parameters. If we consider Example 2.4.2, the leaks and tank volumes are always positive. Failure to incorporate this information in the estimator, as illustrated in Examples 2.4.1 and 2.4.2, may lead to poor estimates.

Asymmetric distributions By piecing together truncated distributions, it is possible to generate asymmetric distributions. The need for asymmetric distributions is illustrated in Example 2.4.2, where mass enters the equalizing tank at a different frequency and magnitude than it would leave at. The inability to model this behavior may lead to spurious estimates as illustrated by the Kalman filter's low estimate of the total losses due to the leak.

Model Simplification Whereas truncated and asymmetric distributions only alter the description of the unknown disturbances, state constraints alter the probabilistic structure of the estimation problem by correlating the disturbances with the state. The advantage is that one can use the correlations to simplify the model significantly. This idea is illustrated in Example 2.4.3 where

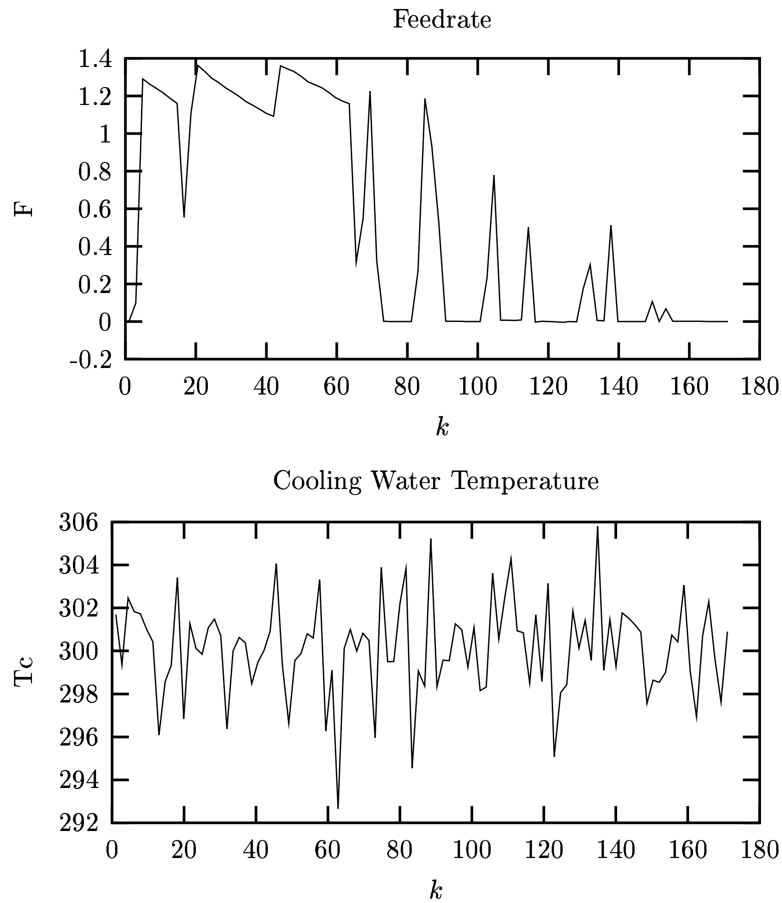


Figure 2.15: Reactor inputs.

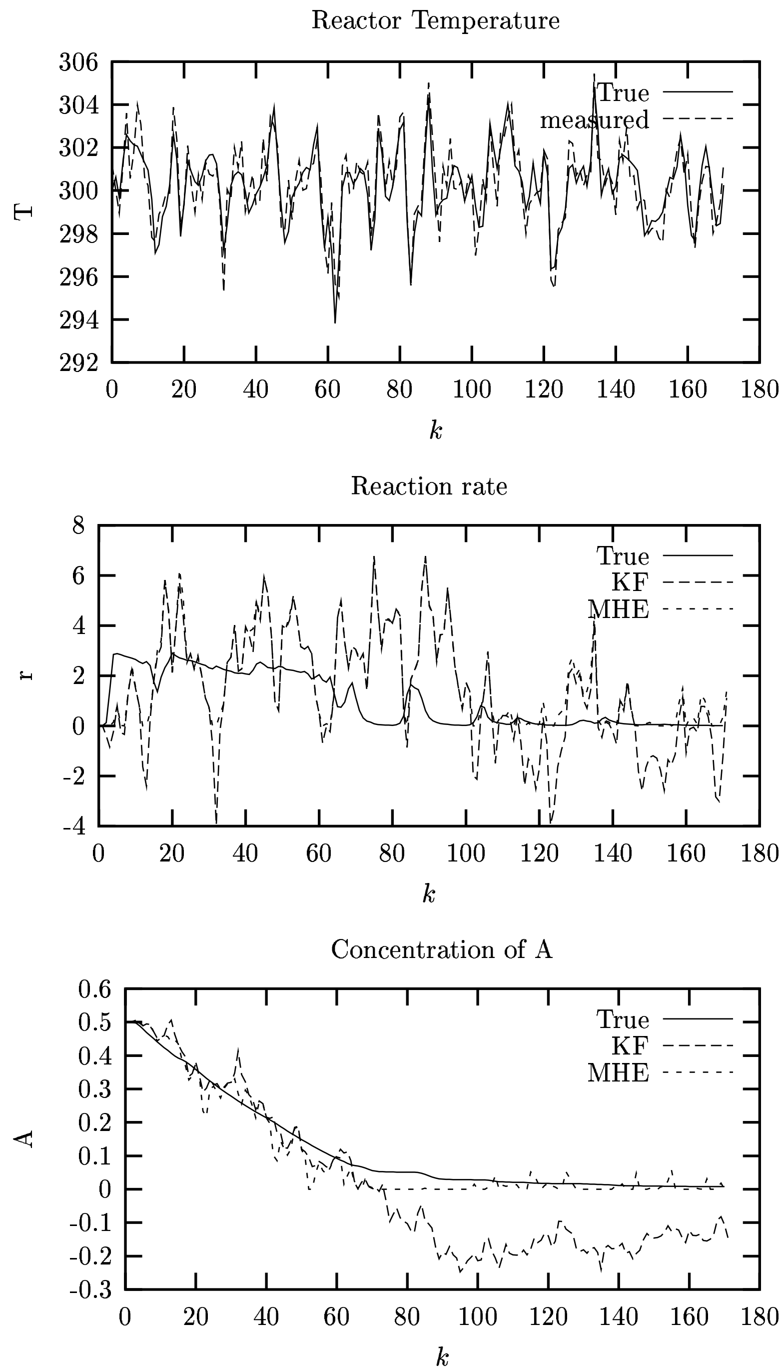


Figure 2.16: Comparison of Estimates

a simplified model of the semibatch reactor using reaction calorimetry coupled with constraints allows for accurate concentration estimates.

Reconciling Conservation Laws Poor measurements may lead to estimates that violate the conservation laws used to model the system. As one often expects the estimates to satisfy the conservation laws, direct enforcement may require inequality constraints. In Example 2.4.3, the estimates of the reaction rate are too high, and the estimates need to be adjusted in order to prevent negative concentration estimates. From a numerical perspective, one may use constraints to prevent the optimization algorithm from choosing spurious iterates that lead to computational problems regarding the solution of the conservation laws and the associated constitutive relations.

The strength and weakness of MHE is the use of mathematical programming. For reasonable models, the optimization problems can be solved in a few seconds on desktop computers using standard software. However, for some problems this performance is insufficient. With the increasing power of computers and improved algorithms (i.e. algorithms now solve quadratic programs in polynomial time), MHE will become an alternative for an expanding class of constrained process monitoring problems in the near future.

Chapter 3

Constrained State Estimation for Nonlinear Discrete-Time Systems[1]

3.1 Introduction

In this chapter we investigate online optimization strategies for estimating the state of systems modeled by a nonlinear difference equation of the form

$$\begin{aligned} x_{k+1} &= f_k(x_k, w_k) \\ y_k &= h_k(x_k) + v_k, \end{aligned} \tag{3.1}$$

where it is known that the states and disturbances satisfy the following constraints

$$x_k \in \mathbb{X}_k, \qquad w_k \in \mathbb{W}_k, \qquad v_k \in \mathbb{V}_k.$$

We assume, for all $k \geq 0$, the functions $f_k : \mathbb{X}_k \times \mathbb{W}_k \to \mathbb{X}_k$ and $h_k : \mathbb{X}_k \to \mathbb{R}^p$ and the sets $\mathbb{X}_k \subseteq \mathbb{R}^n$, $\mathbb{W}_k \subseteq \mathbb{R}^m$, and $\mathbb{V}_k \subseteq \mathbb{R}^p$ are closed with $0 \in \mathbb{W}_k$ and $0 \in \mathbb{V}_k$.

Let $x(k; z, l, \{w_j\})$ denote the solution of the difference equation (3.1) at time k when the initial state is z at time l and the input disturbance sequences is $\{w_j\}_{j=l}^{k}$. When we consider the disturbance free response of the system, i.e. $\{w_j\} = \{0\}$, we use the following notational simplification $x(k; z, l)$. Let $y(k; z, l, \{w_j\}) := h_k(x(k; z, l, \{w_j\}))$ denote the output response of the difference equation (3.1) at time k when the initial state is z at time l and the input disturbance sequences is $\{w_j\}_{j=l}^{k}$. We use the notational simplification $y(k; z, l) := h_k(x(k; z, l))$ for the disturbance free output response of the system. Note the difference between y_k and $y(k; z, l, \{w_j\})$. The vector y_k denotes the observed output at time k and the vector $y(k; z, l, \{w_j\})$ denotes the predicted output at time k when the initial condition at time l is z and the disturbance sequence is $\{w_j\}_{j=l}^{k}$.

We formulate the constrained estimation problem, for $T \geq 0$, as the solution to the following optimal control problem

$$P_1(T) : \qquad \Phi_T^* = \min_{x_0, \{w_k\}_{k=0}^{T-1}} \{\Phi_T(x_0, \{w_k\}) \ : \ (x_0, \{w_k\}) \in \Omega_T\}$$

where the objective function is defined by

$$\Phi_T(x_0, \{w_k\}) := \sum_{k=0}^{T-1} L_k(w_k, v_k) + \Gamma(x_0),$$

[1]**Portions of this chapter were published in Rao and Rawlings (1998*a*) and Rao, Rawlings and Mayne (2000)**

the constraint set is defined by

$$\Omega_T := \left\{ (x_0, \{w_k\}) \ : \ \begin{array}{l} x(k; x_0, 0, \{w_j\}) \in \mathbb{X}_k, \quad k = 0, \dots T, \\ w_k \in \mathbb{W}_k, \quad k = 0, \dots, (T-1), \\ v_k = y_k - y(k; x_0, 0, \{w_j\}) \in \mathbb{V}_k, \quad k = 0, \dots, (T-1) \end{array} \right\},$$

and $v_k := y_k - y(k; x_0, 0, \{w_j\})$. We assume the stage cost function $L_k : \mathbb{W}_k \times \mathbb{V}_k \to \mathbb{R}$ for all $k \geq 0$ and the initial penalty $\Gamma : \mathbb{X}_0 \to \mathbb{R}$. The initial penalty $\Gamma(\cdot)$ summarizes the prior information at time $k = 0$ and satisfies $\Gamma(\hat{x}_0) = 0$, where $\hat{x}_0 \in X_0$ is the *a priori* most likely value of x_0, and $\Gamma(x) > 0$ for $x \neq \hat{x}_0$; The initial penalty $\Gamma(\cdot)$ is part of the data of the state estimation problem. Typically

$$\Gamma(x) := (x - \hat{x}_0)^T \overline{\Pi}_0^{-1} (x - \hat{x}_0),$$

where the matrix $\overline{\Pi}$ is symmetric positive definite. In this case, the given data $(\hat{x}_0, \overline{\Pi})$ determines $\Gamma(\cdot)$. The solution to $P_1(T)$ at time T is the pair

$$\left(\hat{x}_{0|T-1}, \{\hat{w}_{k|T-1}\}_{k=0}^{T-1}\right),$$

and that optimal pair yields an estimate $\{\hat{x}_{k|T-1}\}_{k=0}^{T}$ of the actual sequence $\{x_k\}$; the sequence $\{\hat{x}_{k|T-1}\}_{k=0}^{T}$ is the solution of (3.1) with the initial state $\hat{x}_{0|T-1}$ at time $k = 0$ and disturbance sequence $\{\hat{w}_{k|T-1}\}_{k=0}^{T-1}$, i.e.

$$\hat{x}_{k|T-1} := x(k; \hat{x}_{0|T-1}, 0, \{\hat{w}_{j|T-1}\}).$$

To simplify notation $\hat{x}_j := \hat{x}_{j|j-1}$, where $\hat{x}_{0|-1} = \hat{x}_0$.

We refer to the formulation $P_1(T)$ as the full information problem and $\hat{x}_k$ as the full information estimate of x_k, because all of the available information $\{y_k\}_{k=0}^{T-1}$ is processed. The problem $P_1(T)$ has T stages, so the computational complexity scales at least linearly with T. Unless the process is linear, unconstrained, and the cost functions are quadratic, in which case the optimal estimator is the Kalman filter and the solution is obtained recursively, the online solution of $P_1(T)$ is impractical, because the computational burden increases with time. To make the problem tractable, we need to bound the problem size. One strategy to reduce $P_1(T)$ to a fixed-dimension optimal control problem is to employ a moving horizon approximation. Unlike the full information problem, MHE does not estimate the full state sequence $\{x_k\}_{k=0}^{T}$. Rather, MHE estimates the truncated sequence $\{x_k\}_{k=T-N}^{T}$. The key to preserving stability and performance is how one approximately summarizes the past data.

Consider again the problem $P_1(T)$. We can arrange the objective function by breaking the time interval into two pieces as follows.

$$\Phi_T(x_0, \{w_k\}) = \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \sum_{k=0}^{T-N-1} L_k(w_k, v_k) + \Gamma(x_0).$$

Because we use a state-variable description of the system, the quantity

$$\sum_{k=T-N}^{T-1} L_k(w_k, v_k)$$

depends only on the state x_{T-N} and the sequences $\{w_k, v_k\}_{k=T-N}^{T-1}$. Exploiting the relation using forward dynamic programming, we can establish the equivalence between a full information problem and an estimation problem with a fixed-size estimation window.

Consider the reachable set of states at time τ generated by a feasible initial condition x_0 and disturbance sequence $\{w_k\}_{k=0}^{\tau-1}$:

$$\mathcal{R}_\tau = \{x(\tau; x_0, 0, \{w_j\}) \ : \ (x_0, \{w_j\}) \in \Omega_\tau\}$$

We define the **arrival cost** [2] at time τ and for the state $z \in \mathcal{R}_\tau$ as

$$\mathcal{Z}_\tau(z) := \min_{x_0, \{w_k\}_{k=0}^{\tau-1}} \left\{ \Phi_\tau(x_0, \{w_k\}) \; : \; (x_0, \{w_k\}) \in \Omega_T, \; x(\tau; x_0, 0, \{w_j\}) = z \right\}.$$

It follows that $\mathcal{Z}_0(\cdot) = \Gamma(\cdot)$. Arrival cost is a fundamental concept in MHE, because we can reformulate $P_1(T)$, for $T > N$, as the following equivalent optimal control problem

$$P_1'(T) \qquad \Phi_T^* = \min_{z, \{w_k\}_{k=T-N}^{T-1}} \left\{ \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \mathcal{Z}_{T-N}(z) \; : \; (z, \{w_k\}) \in \Omega_T^N \right\},$$

where the constraint set is defined by

$$\Omega_T^N := \left\{ (z, \{w_k\}) \; : \; \begin{array}{l} x(k; z, T-N, \{w_j\}) \in \mathbb{X}_k, \;\; k = (T-N), \ldots, T, \\ w_k \in \mathbb{W}_k, \;\; k = (T-N), \ldots, (T-1), \\ v_k = y_k - y(k; z, T-N, \{w_j\}) \in \mathbb{V}_k, \\ \qquad k = (T-N), \ldots, (T-1) \end{array} \right\}$$

and $v_k := y_k - y(k; z, T-N, \{w_k\})$. When $T \leq N$, the optimal control problem $P_1'(T)$ is defined to be $P_1(T)$. It is relatively straightforward to demonstrate the equivalence of the solutions to $P_1(T)$ and $P_1'(T)$ using forward dynamic programming (c.f. (Bersekas 1995*a*)).

Optimality guarantees $\mathcal{Z}_T(z) \geq \Phi_T^*$ for all $z \in \mathcal{R}_T$ and $\mathcal{Z}_T(\hat{x}_T) = \Phi_T^*$. We can view, therefore, the arrival cost as an equivalent statistic (Striebel 1965) for summarizing the past data $\{y_k\}_{k=0}^{T-N-1}$ not explicitly accounted for in the objective function of $P_1'(T)$. The arrival cost serves as an equivalent statistic by penalizing the deviation of x_{T-N} away from $\hat{x}_{T-N}$. If we have high (low) confidence in the optimal estimate $\hat{x}_{T-N}$, then the cost of choosing x_{T-N} far away from $\hat{x}_{T-N}$ is large (small).

For the vast majority of systems, we do not possess an algebraic expression for the arrival cost. Notable exceptions are unconstrained linear systems with quadratic objectives, where the estimate $\hat{x}_j$ is now the standard Kalman estimate of the state x_j. Assume the functions $f_k(\cdot)$ and $h_k(\cdot)$ are defined by

$$f_k(x_k) := A_k x + G_k w, \qquad h_k(x) := C_k x,$$

and the stage penalties $L_k(\cdot)$ are defined by

$$L_k(w, v) := w^T Q_k^{-1} w + v^T R_k^{-1} v,$$

where the matrices Q_k and R_k are symmetric positive definite. For this case, the initial penalty is defined as

$$\Gamma(x) := (x - \hat{x}_0)^T \overline{\Pi}_0^{-1} (x - \hat{x}_0),$$

and the arrival cost is given by

$$\mathcal{Z}_j(z) := (z - \hat{x}_j)^T \Pi_j^{-1} (z - \hat{x}_j) + \Phi_j^* \tag{3.2}$$

assuming the matrix Π_j is invertible. The matrix sequence $\{\Pi_j\}$ is obtained by solving the matrix Riccati equation

$$\Pi_{j+1} = G_j Q_j G_j^T + A_j \Pi_j A_j^T - A_j \Pi_j C_j^T (R_j + C_j \Pi_j C_j^T)^{-1} C_j \Pi_j A_j^T. \tag{3.3}$$

[2]Other researchers have used the term **cost to come** (c.f (Başar and Bernhard 1995)) or **cost to arrive** (c.f. (Verdu and Poor 1987)).

with the initial condition $\Pi_0 = \overline{\Pi}_0$. One obtains this result by deriving the deterministic Kalman filter using forward dynamic programming (c.f. Cox (1964)).

When the system is nonlinear or constrained, an algebraic expression for the arrival cost does not exist, yet we require one to implement successfully the estimator. Ideally, we want the moving horizon estimate as close as possible to the full information estimate. One solution is to formulate MHE as the solution to a numerically tractable though approximate version of $P_1'(T)$. An approximation $\hat{\mathcal{Z}}_j(\cdot)$ of the arrival cost $Z_j(\cdot)$ may be used to account for the data not included in the estimation window. The past data are accounted for approximately with our choice of $\hat{\mathcal{Z}}_j(\cdot)$ by penalizing deviation away from the past estimate $\hat{x}_j$ in accordance with our confidence in the estimate. Because this choice is an approximation, we need to ensure that MHE does not improperly weight the old data. Estimator divergence may result if the approximation biases the past data by weighting the past estimates too strongly, while performance may suffer if the approximation insufficiently weights the past data. In Section 3.4 we discuss the stability implications of approximate representations of the arrival cost.

We formulate, for $T > N$, the moving horizon approximation to the full information estimation problem, or MHE, as the following optimal control problem

$$P_2(T) \qquad \hat{\phi}_T = \min_{z,\{w_k\}_{k=T-N}^{T-1}} \left\{ \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \hat{\mathcal{Z}}_{T-N}(z) \ : \ (z, \{w_k\}) \in \Omega_T^N \right\}$$

where $v_k := y_k - y(k; z, T-N, \{w_j\})$ and $\hat{\mathcal{Z}}_j : \mathbb{X}_j \to \mathbb{R}$ for all $j \geq 0$. The moving horizon cost $\hat{\phi}_T$ is an approximation of Φ_T^* obtained by replacing the (uncomputable) arrival cost $\mathcal{Z}_{T-N}(\cdot)$ with an approximation $\hat{\mathcal{Z}}_{T-N}(\cdot)$. We choose $\hat{\mathcal{Z}}_0(\cdot) = \Gamma(\cdot)$. When $T \leq N$, the optimal control problem $P_2(T)$ is defined to be $P_1(T)$. The solution to $P_2(T)$ at time T is the pair $(z^*, \{\hat{w}_{k|T-1}^{\text{mh}}\}_{k=T-N}^{T-1})$, which, when used as data in the system equation (3.1), yields $\{\hat{x}_{k|T-1}^{\text{mh}}\}_{k=T-N}^{T}$, i.e.

$$\hat{x}_{k|T-1}^{\text{mh}} := x(k; z^*, T-N, \{\hat{w}_{j|T-1}^{\text{mh}}\}).$$

For simplicity $\hat{x}_j^{\text{mh}} := \hat{x}_{j|j-1}^{\text{mh}}$, where $\hat{x}_{0|-1}^{\text{mh}} = \hat{x}_0$.

One strategy to approximate the arrival cost $\mathcal{Z}_T(\cdot)$ is to employ a first order Taylor series approximation of the model (3.1) around the estimated trajectory $\{\hat{x}_k^{\text{mh}}\}_{k=0}^{T}$. This strategy yields an extended Kalman filter covariance update formula for constructing $\hat{\mathcal{Z}}_T(\cdot)$. We interpret this strategy as a neighboring extremal paths strategy in the context of estimation. Neighboring extremal paths are used to generate approximate optimal feedback laws for nonlinear systems by employing extended linearization (Bryson and Ho 1975). The basic idea is as follows. If the deviation from the optimal path is small, then a linear approximation at the optimal path accurately describes the neighboring path.

Suppose the model functions $f_k(\cdot)$ and $h_k(\cdot)$ and the cost functions $L_k(\cdot)$ are sufficiently smooth and

$$\Gamma(x) := (x - \hat{x}_0)^T \overline{\Pi}_0^{-1} (x - \hat{x}_0).$$

Let

$$A_k := \left. \frac{\partial f_k(x, 0)}{\partial x} \right|_{\hat{x}_k^{\text{mh}}}, \qquad G_k := \left. \frac{\partial f_k(\hat{x}_k^{\text{mh}}, w)}{\partial w} \right|_{w=0}, \qquad C_k := \left. \frac{\partial h_k(x)}{\partial x} \right|_{\hat{x}_k^{\text{mh}}},$$

denote the linearized dynamics of the system (3.1) and

$$R_k^{-1} := \left. \frac{\partial^2 \underline{L}_k(0, v)}{\partial v \partial v^T} \right|_{\hat{x}_k^{\text{mh}}}, \qquad N_k := \left. \frac{\partial^2 \underline{L}_k(w, v)}{\partial w \partial v^T} \right|_{w=0,\ \hat{x}_k^{\text{mh}}},$$

$$Q_k^{-1} := \left. \frac{\partial^2 \underline{L}_k(w, v)}{\partial w \partial w^T} \right|_{w=0,\ \hat{x}_k^{\text{mh}}},$$

denote the linearized stage penalties $L_k(\cdot)$, then, if we assume for simplicity $N_k = 0$, we approximate the arrival cost as

$$\hat{\mathcal{Z}}_T(z) = (z - \hat{x}_T^{\text{mh}})^T \Pi_T^{-1} (z - \hat{x}_T^{\text{mh}}) + \hat{\phi}_T,$$

assuming the matrix Π_T is invertible, where the matrix sequence $\{\Pi_j\}$ is obtained by solving the matrix Riccati equation (3.3) subject to the initial condition $\Pi_0 = \overline{\Pi}_0$. This result is equivalent to the covariance update formula for the extended Kalman filter. See Jazwinski (1970) for further details.

The remaining chapter is organized as follows. Section 3.2 introduces the notation, definitions, and basic assumptions necessary for establishing stability. We establish sufficient conditions for the stability of the full information estimator in Section 3.3. In Section 3.4, we extent the results of Section 3.3 to derive sufficient condition for the stability of MHE. We also propose a prototype algorithm for MHE. Obtaining global solutions to nonlinear optimal control problems presents a formidable barrier to online implementation. In Section 3.5, we investigate suboptimal strategies that guarantee stability, but do not require global solutions. In Section 3.6, we examine the dual relationship of MHE and receding horizon control.

3.2 Notation, Definitions, and Basic Assumptions

The Cartesian product $\times_{k=1}^{N} \mathbb{A}$ of a set $\mathbb{A}$ is denoted by $\mathbb{A}^N$. We use the symbol $\|\cdot\|$ to denote any vector norm in $\mathbb{R}^n$ (where the dimension n follows from context). Let $\mathbb{R}_{\geq 0}$ denote the nonnegative real numbers, $\underline{\mathcal{C}}(\mathbb{R}^n)$ denote the space of lower semi-continuous functions that map from $\mathbb{R}^n$ to $\mathbb{R}$, and $l_2(\mathbb{R}^n)$ denote the space of all sequences $\{a_k\}$ in $\mathbb{R}^n$ for which $\sum_{k=0}^{\infty} |a_k|^2 < \infty$, where $|x| = \sqrt{x^T x}$. For $\epsilon > 0$, $B_\epsilon := \{x \ : \ \|x\| \leq \epsilon\}$. Let $a \vee b := \max\{a, b\}$. For notational simplicity, we make that following definition: $\hat{(\cdot)}_k := \hat{(\cdot)}_{k|k-1}$.

Definition 3.2.1 *A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a* **K-function** *if it is continuous, strictly monotone increasing, $\alpha(x) > 0$ for $x \neq 0$, $\alpha(0) = 0$, and $\lim_{x\to\infty} \alpha(x) = \infty$.*

Throughout the paper we use the following elementary properties of K-functions.

Fact 3.2.2 *Suppose $\alpha(\cdot)$ is a K-function. Then, the function $\alpha(\cdot)$ and its inverse $\alpha^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ are continuous (Royden 1988). Furthermore, $\alpha^{-1}(\cdot)$ is a K-function.*

Fact 3.2.3 *The space of K-functions is closed under addition, composition, and positive scalar multiplication. For example, if $\alpha(\cdot)$ and $\beta(\cdot)$ are K-functions, then $\alpha \circ \beta(\cdot)$, $\alpha(\cdot) + \beta(\cdot)$, $c\alpha(\cdot)$ for $c > 0$ are K-functions.*

Definition 3.2.4 *A system is* **uniformly observable** *if there exists a positive integer N_o and a K-function $\varphi(\cdot)$ such that for any two states x_1 and x_2,*

$$\varphi(\|x_1 - x_2\|) \leq \sum_{j=0}^{N_o - 1} \|y(k+j; x_1, k) - y(k+j; x_2, k)\|,$$

for all $k \geq 0$.

The observability condition states that if the prediction residuals are small, then the estimation error is small. Mathematically this condition requires that the mapping from the state x to the sequence

$\{y(k+j;x,k)\}_{j=0}^{N_o-1}$ is bounded and one-to-one. When the system is linear (i.e. $f_k(x) = A_k x$ and $h_k(x) = C_k x$), then the uniform observability condition is satisfied when the observability Grammian

$$V_k := \sum_{j=0}^{N-1} A_{k+j}^{jT} C_{k+j}^T C_{k+j} A_{k+j}^j$$

is positive definite for all $N \geq n$ and $k \geq 0$.

In order to guarantee the problems $P_1(T)$ and $P_2(T)$ are well posed, we require that the model (3.1), stage cost functions $L_k(\cdot)$, and initial penalty $\Gamma(\cdot)$ satisfy the following conditions.

A0 The functions $f_k(\cdot)$ and $h_k(\cdot)$ are globally Lipschitz continuous with constants c_f and c_h for all $k \geq 0$.

A1 $L_k(\cdot) \in \underline{\mathcal{C}}(\mathbb{W}_k \times \mathbb{V}_k)$ for all $k \geq 0$ and $\Gamma(\cdot) \in \underline{\mathcal{C}}(\mathbb{X}_0)$.

A2 There exist K-functions $\eta(\cdot)$ and $\gamma(\cdot)$ such that

$$\begin{aligned} \eta(\|(w,v)\|) \leq L_k(w,v) \leq \gamma(\|(w,v)\|), \\ \eta(\|x-\hat{x}_0\|) \leq \Gamma(x) \leq \gamma(\|x-\hat{x}_0\|), \end{aligned}$$

for all $(w,v) \in (\mathbb{W}_k \times \mathbb{V}_k)$, $x \in \mathbb{X}_0$, $\hat{x}_0 \in \mathbb{X}_0$, and $k \geq 0$.

We need also to impose similar conditions on the approximate arrival cost $\hat{\mathcal{Z}}_k(\cdot)$. However, unlike the initial penalty, the minimal value of the arrival cost is greater than zero (recall $\mathcal{Z}_k(x) \geq \hat{\phi}_k$ for all $x \in \mathbb{X}_k$ with $\mathcal{Z}_k(\hat{x}_k) = \hat{\phi}_k$) and the approximate arrival $\hat{\mathcal{Z}}_k(x)$ may not be bounded below by $\|x\|$ for reasons that become apparent in Section 3.4. We require instead $\hat{\mathcal{Z}}_k(\cdot)$ satisfies the following condition.

C1 There exist K-function $\bar{\gamma}(\cdot)$ such that

$$0 \leq \hat{\mathcal{Z}}_k(z) - \hat{\phi}_k \leq \bar{\gamma}(\|z - \hat{x}_k^{\mathrm{mh}}\|)$$

for all $z \in \mathbb{X}_T$, and $T \geq 0$.

3.2.1 Observer Stability

The following discussion of observer stability is premised on classical Lyapunov stability theory for dynamical systems. The concepts are completely analogous to their classical counterpart. To account for constraints, we have modified the definition of stability in an analogous manner to Keerthi and Gilbert (1988).

Before proceeding with the abstract definition of observer stability, we first introduce the concept of an observer. When the state is not directly available from the measurements, it is necessary to infer the current state of the system. The observer problems considers inferring the current state of the following deterministic (i.e. $(w_k, v_k) \equiv 0$) system

$$x_{k+1} = f_k(x_k), \tag{3.4a}$$

$$y_k = h_k(x_k). \tag{3.4b}$$

When the functions $f_k(\cdot)$ and $h_k(\cdot)$ are linear and time invariant, the Luenberger observer (Luenberger 1966, Luenberger 1971) solves the observer problem. The basic strategy of the Luenberger observer is to design a linear output feedback law that stabilizes the error dynamics of the system (3.4). Note the observer problem is solved if the initial condition x_0 is known exactly, because the model will track the system perfectly.

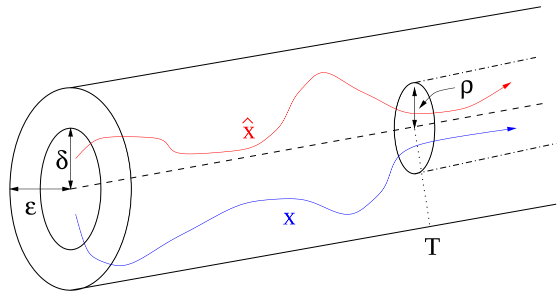


Figure 3.1: Definition of an asymptotically stable observer.

One strategy for the observer problem is to solve directly for x_0. If the system (3.4) is uniformly observable and there is sufficient data, the initial condition can be uniquely determined from the data. This strategy was employed iteratively using moving horizon observers by Jang et al. (1986), Zimmer (1994), Moraal and Grizzle (1995), and Michalska and Mayne (1995). The terminology moving horizon observer is used because only the state of a deterministic system is reconstructed. This strategy is limited, because the problem statement is an idealization: our knowledge of the true system is not complete. In addition, state and measurement noise prevent a consistent state estimate being recovered from a fixed window of data. By assuming a deterministic model structure, the moving horizon observer strategies may amplify rather than filter the effect of noise on the estimate. Finally, the iterative strategies may break down, because cost decreases are not always possible. Even though the problem statement is an idealization, it is prudent to ask whether the estimation strategy proposed in this work can address this *simple* problem. The ability to reconstruct the state of the system (3.4) from the process measurements is what we define as observer stability.

We begin our analysis by making precise the notion of observer stability. The concept of observer stability and asymptotic observer stability is equivalent to stating that the estimation error $e_k := \|x(k; x_0, 0) - \hat{x}_k\|$, where x_0 is the initial state of the system (3.4), is bounded by $\|e_0\|$ and converges to 0 as $k \to \infty$.

Definition 3.2.5 *An estimator is an* **asymptotically stable observer** *for the system (3.4) if, for every initial condition $x_0 \in \mathbb{X}_0$ and every $\epsilon > 0$, there corresponds a number $\delta > 0$ and a positive integer $\bar{T}$ such that if $\|x_0 - \hat{x}_0\| \leq \delta$ and $\hat{x}_0 \in \mathbb{X}_0$, then $\|x(T; x_0, 0) - \hat{x}_T\| \leq \epsilon$ for all $T \geq \bar{T}$. Furthermore, for all $x_0 \in \mathbb{X}_0$, $\hat{x}_T \to x(T; x_0, 0)$ as $T \to \infty$.*

3.2.2 Feasibility and Constraints

To guarantee that a solution exists to either $P_1(T)$ or $P_2(T)$, we require that the feasible region is nonempty. The implications of constraints are more subtle for an estimator than for a regulator. In particular, an estimator has no control over the evolution of the state of the system (3.4). A stable observer may not exist when the constraints are poorly chosen. For example, consider the system with $f_k(\cdot) = x_k + w_k$, $h_k(\cdot) = x_k$, and the initial condition $x_0 = 0$. If we choose the constraints such that $\mathbb{X}_k = \{x : |x| \geq 1\}$ for $k \geq 0$, then there does not exist a stable observer that is able to reconstruct the state of the system. Furthermore, as the constraint $\mathbb{V}_k$ directly relates feasibility to the outputs $\{y_k\}_{k=0}^{T-1}$, we need, therefore, to characterize the conditions of the system generating the output measurements. The focus of the chapter is on obtaining sufficient conditions for stability of moving horizon observers, so we limit ourselves in the following assumption to the case when the system (3.4) generates the data.

In addition to assuming the feasible region is nonempty, in order to guarantee a stable (constrained) observer exits, we need to assume also that the infinite-time cost is bounded by the initial estimation error $\|x_0 - \hat{x}_0\|$.

A3 Suppose the system (3.4) with initial condition x_0 generates the data (i.e. $y_k = y(k; x_0, 0)$). There exists an initial condition $x_{0|\infty}$, disturbance sequence $\{w_{k|\infty}\}_{k=0}^{\infty}$, and a K-function $\sigma(\cdot)$ such that, for all $\hat{x}_0 \in \mathbb{X}_0$,

$$\lim_{T\to\infty} \Phi_T(x_{0|\infty}, \{w_{k|\infty}\}) \leq \sigma(\|x_0 - \hat{x}_0\|),$$

where, for all $k \geq 0$, $\big(x_{0|\infty}, \{w_{j|\infty}\}\big) \in \Omega_k$.

If we suppose the system (3.4) with the initial condition x_0 satisfies the constraints $x(k; x_0, 0) \in \mathbb{X}_k$ for all $k \geq 0$, then we can satisfy assumption **A3** if we choose $x_{0|\infty} = x_0$ and $w_{k|\infty} = 0$. Recall $0 \in \mathbb{W}_k$ and $0 \in \mathbb{V}_k$. If we choose $\sigma(\cdot) = \gamma(\cdot)$, where the K-function $\gamma(\cdot)$ is defined in **A2**, then we obtain the stated conditions of assumption **A3**. In particular, for $\hat{x}_0 \in \mathbb{X}_0$,

$$\lim_{T\to\infty} \Phi_T(x_{0|\infty}, \{w_{k|\infty}\}) = \Gamma(x_0) \leq \gamma(\|x_0 - \hat{x}_0\|).$$

Indeed, the system (3.4) should satisfy the states constraint for all $x_0 \in \mathbb{X}_0$. Constraints in estimation should reflect additional insight about the evolution of the system. If the system does not obey the constraints, then no additional insight is offered by constraining the state estimator, and one may, in fact, adversely affect estimator performance.

When we consider a constrained estimator in practice, assumption **A3** is, aside from theoretical considerations, meaningless. We do not expect either the system (3.4) or even the system (3.1) generates the data $\{y_k\}$. Poorly designed constraints, therefore, may cause the constrained estimator to fail due to infeasibility or, at the very least, perform poorly. Characterizing the set $\mathbb{W}_k$ is relatively straightforward, though one may experience problems if one improperly characterizes the sets $\mathbb{V}_k$, due to the possibility of outliers, and $\mathbb{X}_k$, for reasons discussed next. One can guarantee a solution exists to either $P_1(T)$ or $P_2(T)$, without explicit reference to the system (3.4), if the model (3.1) and constraints satisfy the following condition[3]

B1 For all $T \geq 0$, there exists $\big(x_0^1, \{w_k^1\}\big) \in \Omega_T$.

The condition **B1** is nonstandard; one usually chooses an exact model of the plant and, separately, the characteristics of the disturbances, such as boundedness, or that the disturbances are independent and identically distributed with known (zero) mean and variance. The properties of the model and disturbances are distinct. Assumption **B1**, on the other hand, implicitly states that the model is in error, because the disturbance free evolution of the system (3.1), i.e. $x(k; x_0, 0)$, may not automatically satisfy the state constraints for some $x_0 \in \mathbb{X}_0$. Enforcing the state constraints may require a non-zero disturbance sequence $\{w_k\}$, thus implicitly using $\{w_k\}$ to account for model error. Furthermore, state constraints may correlate the disturbances $\{w_k\}$. If we consider, for example, the system $x_{k+1} = x_k + w_k$ subject to the constraints $x_k \geq 0$ and $w_k \leq 0$, then a large (negative) w_0 implies the future $\{w_k\}$ are small. Alternatively, the disturbance w_k is correlated with the state x_k. If x_k is small, then w_k needs to be small in order to satisfy the state constraint.

Suppose instead the constraints $\mathbb{X}_k$ denote the subset of $\mathbb{R}^n$ for which the model (3.1) is defined. In particular, the sets $\mathbb{X}_k$ define the flow of the difference equation (3.1) for a subset of initial conditions $\mathbb{X}_0$.

[3] The conditions of the existence results, Propositions 3.3.1 and 3.4.3, are satisfied by design if assumption **A3** is replaced with either **B1** or **B2**. Existence is established if the pair $x_{0|\infty}$ and $\{w_{k|\infty}\}$ is replaced with the pair x_0^1 and $\{w_k^1\}$

B2 For all $T \geq 0$ and $x_0^1 \in \mathbb{X}_0$, $\left(x_0^1, \{w_k^1\}\right) \in \Omega_T$ for all disturbances $\{w_k^1\}_{k=0}^{T-1}$ satisfying the constraints $w_k^1 \in \mathbb{W}_k$.

Condition **B2**, obviously, implies **B1**. Furthermore, the "statistics" of the disturbances w_k are decoupled from the state of the system x_k. Our decision to include **B1** is practically motivated. One is typically unable, or unwilling, to model separately the system and disturbances; the modeling effort is too great, and the disturbances w_k and v_k are convenient to account for model uncertainty. State constraints may then be necessary to complete the model (e.g. complete the conservation laws). The advantage of state constraints is that they allow for simplified models (c.f. (M'hamdi et al. 1996)). While not always theoretically satisfying, state constraints may be practically appealing. The issues regarding state constraints have not been resolved completely, and the practitioner should be cognizant of the differences between, and possible implications of, **B1** and **B2**.

3.3 Full Information Estimation

Full information implies that at time T the data $\{y_k\}_{k=0}^{T-1}$ are employed; the complexity of the estimation problem, therefore, increases with time T. However, there is no approximation error in contrast to the moving horizon estimator where an estimate $\hat{\mathcal{Z}}_{T-N}(\cdot)$ of the arrival cost $\mathcal{Z}_{T-N}(\cdot)$ is employed. Hence, we consider first the simpler full information case. We begin by providing sufficient conditions for the existence of a solution to $P_1(T)$. We then prove stability.

Proposition 3.3.1 *If assumptions* **A0**–**A3** *hold, then a solution exists to $P_1(T)$ for all $\hat{x}_0 \in \mathbb{X}_0$ and $T \geq 0$.*

Proof. By assumption **A3**, the feasible region is non-empty for all T. Let $\Phi_T^1 = \Phi_T(x_{0|\infty}, \{w_{k|\infty}\})$ denote the finite cost, by **A2** , associated with $x_{0|\infty}$ and the sequence $\{w_{k|\infty}\}_{k=0}^{T-1}$. Consider the set

$$\Lambda = \left\{ \left(z, \{w_k\}_{k=0}^{T-1}\right) \ : \ (z, \{w_k\}) \in \Omega_T, \ \ \Phi_T(z, \{w_k\}) \leq \Phi_T^1 \right\}.$$

By assumption **A0** and **A1**, the function $\Phi_T(\cdot)$ is lower semi-continuous. The set

$$F = \left\{ \left(z, \{w_k\}_{k=0}^{T-1}\right) \ : \ \Phi_T(z, \{w_k\}) \in [0, \Phi_T^1] \right\}$$

is closed, because the inverse image of a closed set under a lower semi-continuous function is closed (Berge 1963). The set Ω_T is closed, because the functions $f_k(\cdot)$ and $h_k(\cdot)$ are continuous and the sets $\mathbb{X}_k$, $\mathbb{W}_k$, and $\mathbb{V}_k$ are closed for all $k \geq 0$. The set Λ is closed, because it is the intersection of the closed sets F and Ω_T. By assumption **A2**, there exists a K-function $\eta(\cdot)$ such that $L_k(w, v) \geq \eta(\|w, v\|)$ and $\Gamma(x) \geq \eta(\|x - \hat{x}_0\|)$. These inequalities imply the set F is also bounded, because $\Gamma(x_0) \leq \Phi_T^1$ and $L_k(w_k, v_k) \leq \Phi_T^1$ for $k = 0, \ldots, T-1$. The set Λ is bounded, because $\Lambda \subseteq F$. Hence, the set Λ is compact. Existence of a solution follows from the Weierstrass Maximum Theorem. □

Before proving stability, we require the following technical lemma.

Lemma 3.3.2 *Let $x_0 \in \mathbb{X}_0$. Suppose* **A0**–**A2** *are true and the system (3.1) is uniformly observable. For all $N \geq N_o$, if*

$$\sum_{k=T-N}^{T-1} L_k(\hat{w}_{k|T-1}, \hat{v}_{k|T-1}) \to 0$$

then $\|x(T; x_0, 0) - \hat{x}_T\| \to 0$ for $T \geq N$.

Proof. The proof is given in Appendix 3.8.1. □

The following proposition establishes stability by demonstrating that the optimal cost function Φ_k^* is nondecreasing and bounded above uniformly for all $k \geq 0$ by the initial estimation error $\|x_0 - \hat{x}_0\|$.

Proposition 3.3.3 *If assumptions* **A0–A3** *hold, and the system (3.1) is uniformly observable, then, for all $\hat{x}_0 \in \mathbb{X}_0$, the full information estimator is an asymptotically stable observer for the system (3.4).*

Proof. We assume throughout the proof $T \geq N_o$ (set $\bar{T} = N_o$). We first demonstrate convergence. Proposition 3.3.1 guarantees a solution exists for all $k \geq 0$. By optimality and **A3** , we have that $\Phi_k^* \leq \sigma(\|x_0 - \hat{x}_0\|)$ for all $k \geq 0$. Recall, x_0 denotes the initial condition of the system (3.4). Writing out the cost function explicitly, we regroup the optimal cost as follows

$$\begin{aligned} \Phi_T^* &= \sum_{k=0}^{T-1} L_k(\hat{w}_{k|T-1}, \hat{v}_{k|T-1}) + \Gamma(\hat{x}_{0|T-1}), \\ &= L_{T-1}(\hat{w}_{T-1|T-1}, \hat{v}_{T-1|T-1}) + \sum_{k=0}^{T-2} L_k(\hat{w}_{k|T-1}, \hat{v}_{k|T-1}) + \Gamma(\hat{x}_{0|T-1}). \end{aligned}$$

Because $\{\hat{w}_{k|T-1}\}_{k=0}^{T-2}$ and $\hat{x}_{0|T-1}$ satisfy the constraints in problem $P_1(T-1)$, optimality implies the following inequality

$$\sum_{k=0}^{T-2} L_k(\hat{w}_{k|T-1}, \hat{v}_{k|T-1}) + \Gamma(\hat{x}_{0|T-1}) \geq \Phi_{T-1}^*.$$

This inequality, in turn, implies

$$\Phi_T^* - \Phi_{T-1}^* \geq L_{T-1}(\hat{w}_{T-1|T-1}, \hat{v}_{T-1|T-1}),$$

so the sequence $\{\Phi_k^*\}$ is monotone nondecreasing. By **A3** and optimality, this sequence is bounded above by $\|x_0 - \hat{x}_0\|$. Therefore, the sequence of optimal costs $\{\Phi_j^*\}$ converges to $\Phi_\infty^* < \infty$, and the partial sum

$$\sum_{k=T-N_o}^{T-1} L_k(\hat{w}_{k|T-1}, \hat{v}_{k|T-1}) \to 0,$$

as $T \to \infty$. By Lemma 3.3.2, the estimation error $\|x(T; x_0, 0) - \hat{x}_T\| \to 0$ as claimed.

To prove stability, let $\epsilon > 0$ and choose $\varrho > 0$ as specified by Lemma 3.3.2 such that if

$$\sum_{k=T-N_o}^{T-1} L_k(\hat{w}_{k|T-1}, \hat{v}_{k|T-1}) \leq \varrho,$$

then $\|x(T; x_0, 0) - \hat{x}_T\| \leq \epsilon$ for all $T \geq N_o$. If we choose $\delta > 0$ such that $\delta < \sigma^{-1}(\varrho)$ (the existence of $\sigma^{-1}(\cdot)$ follows from Fact 3.2.2), then we obtain the following inequality for all $T \geq N_o$:

$$\begin{aligned} \sigma(\delta) &\geq \sum_{k=0}^{T-1} L_k(\hat{w}_{k|T-1}, \hat{v}_{k|T-1}) + \Gamma(\hat{x}_{0|T-1}) \\ &\geq \sum_{k=T-N_o}^{T-1} L_k(\hat{w}_{k|T-1}, \hat{v}_{k|T-1}). \end{aligned}$$

Hence, if the initial estimation error $\|x_0 - \hat{x}_0\| \leq \delta$, then the estimation error

$$\|x(T; x_0, 0) - \hat{x}_T\| \leq \epsilon$$

for all $T \geq N_o$ as claimed. □

3.4 Moving Horizon Estimation

In this section we derive sufficient conditions for the stability of MHE. Our arguments closely follow those used in Section 3.3. We begin by stating conditions on the approximate arrival cost $\hat{\mathcal{Z}}_j(\cdot)$ sufficient to guarantee the stability of MHE. We proceed to derive conditions for the existence of a solution to $P_2(T)$, and we then establish stability. For most nonlinear systems the approximate arrival costs are unable to satisfy *a priori* the stability condition. We conclude the section, therefore, by presenting two prototype algorithms for constrained MHE that relax the stability conditions on the approximate arrival costs.

Ideally the approximate arrival cost $\hat{\mathcal{Z}}_j(\cdot)$ is equal to the arrival cost $\mathcal{Z}_j(\cdot)$. With the notable exception of the unconstrained linear quadratic problem (i.e. the Kalman filter), closed-form expressions for the arrival cost are generally unavailable. To guarantee stability, however, we do not need to construct the arrival cost, but rather require instead that the approximate arrival cost satisfies the following condition.

C2 Let

$$\mathcal{R}_\tau^N = \left\{ x(\tau; z, \tau - N, \{w_k\}) \ : \ (z, \{w_k\}) \in \Omega_\tau^N \right\},$$

where $\mathcal{R}_\tau^N = \mathcal{R}_\tau$ for $\tau \leq N$. For a horizon length N, any time $\tau > N$, and any $p \in \mathcal{R}_\tau^N$, the approximate arrival cost $\hat{\mathcal{Z}}_\tau(\cdot)$ satisfies the inequality

$$\hat{\mathcal{Z}}_\tau(p) \leq \min_{z, \{w_k\}_{k=\tau-N}^{\tau-1}} \left\{ \sum_{k=\tau-N}^{\tau-1} L_k(w_k, v_k) + \hat{\mathcal{Z}}_{\tau-N}(z) \ : \ \begin{array}{l} (z, \{w_k\}) \in \Omega_\tau^N, \\ x(\tau; z, \tau - N, \{w_j\}) = p \end{array} \right\}, \tag{3.5}$$

subject to the initial condition $\hat{\mathcal{Z}}_0(\cdot) = \Gamma(\cdot)$. For $\tau \leq N$, the approximate arrival cost $\hat{\mathcal{Z}}_\tau(\cdot)$ satisfies instead the inequality $\hat{\mathcal{Z}}_\tau(\cdot) \leq \mathcal{Z}_\tau(\cdot)$.

If one views arrival cost as an equivalent statistic for the data, then the inequality (3.5) in condition **C2** states that the approximate arrival cost should not add additional "information" not specified in the data. Loosely speaking, we say a positive function $a(\cdot)$ contains more information than another positive function $b(\cdot)$ if $a(x) \geq b(x)$ for all x of interest. If the inequality (3.5) were strict, then condition **C2** would state there should be some "forgetting" in the estimator.

Remark 3.4.1 *A simple strategy to satisfy condition* **C2** *is to define for time τ the approximate arrival cost as $\hat{\mathcal{Z}}_\tau(\cdot) := \hat{\phi}_\tau$. The inequality (3.5) is satisfied by definition: optimality of $P_2(\tau)$ guarantees that the optimal cost $\hat{\phi}_\tau$ satisfies the inequality (3.5) for all $p \in \mathcal{R}_\tau^N$. This construction was employed by Muske and Rawlings (1995) to generate a stable nonlinear MHE. Without constraints, this choice yields a deadbeat observer.*

Remark 3.4.2 *If we choose*

$$\hat{\mathcal{Z}}_\tau(z) = (z - \hat{x}_\tau^{mh})^T \Pi_\tau^{-1} (z - \hat{x}_\tau^{mh}) + \hat{\phi}_\tau,$$

where the sequence $\{\Pi_j\}$ is obtained by solving the matrix Riccati equation (3.3) subject to the initial condition $\Pi_0 = \overline{\Pi}_0$, then condition **C2** *is satisfied when we consider linear systems with quadratic objectives and convex constraints. The proof of this is given in Chapter 4.*

As a consequence of condition **C1**, motivated by Remark 3.4.1, we cannot invoke the coercivity arguments used in Proposition 3.3.1 to establish the existence of a solution to $P_2(T)$. To guarantee a solution exists, we employ the observability assumption.

Proposition 3.4.3 *If assumptions* **A0**–**A3** *hold, the sequence* $\{\hat{\mathcal{Z}}_j(\cdot)\}$ *satisfies condition* **C1**, *the system (3.4) is uniformly observable, and* $N \geq N_o$, *then a solution exists to* $P_2(T)$ *for all* $\hat{x}_0 \in \mathbb{X}_0$ *and* $T \geq 0$.

Proof. The proof is given in Appendix 3.8.2. □

In the following Proposition we state our fundamental result on MHE. The argument are similar to those used in Proposition 3.3.3. In particular, stability is established by demonstrating that the sequence $\{\hat{\phi}_k\}$ is nondecreasing and bounded above uniformly for $k \geq 0$ by the initial estimation error $\|x_0 - \hat{x}_0\|$.

Proposition 3.4.4 *If assumptions* **A0**–**A3** *hold, the sequence* $\{\hat{\mathcal{Z}}_j(\cdot)\}$ *satisfies the condition* **C1** *and* **C2**, *the system (3.1) is uniformly observable, and* $N \geq N_o$, *then, for all* $\hat{x}_0 \in \mathbb{X}_0$, *MHE is an asymptotically stable observer for the system (3.4).*

Proof. We first prove convergence by demonstrating that $\sigma(\|x_0 - \hat{x}\|)$, where $\sigma(\cdot)$ is defined in **A3**, is a uniform upper bound for $\hat{\phi}_k$. Recall x_0 denotes the initial condition of (3.4). Proposition 3.4.3 guarantees an optimal solution exists for all $k \geq 0$ and $\hat{x}_0 \in \mathbb{X}_0$. Assumption **A2** and condition **C1** guarantee, for all $T \geq N$,

$$\hat{\phi}_T - \hat{\phi}_{T-N} \geq \sum_{k=T-N}^{T-1} L_k(\hat{w}^{\text{mh}}_{k|T-1}, \hat{v}^{\text{mh}}_{k|T-1}). \tag{3.6}$$

We proceed using an induction argument. For $T \leq N$, assumption **A3**, optimality, and condition **C2** imply

$$\begin{aligned}
\sigma(\|x_0 - \hat{x}_0\|) &\geq \sum_{k=0}^{T-1} L_k(w_{k|\infty}, v_{k|\infty}) + \Gamma(x_{0|\infty}), \\
&\geq \mathcal{Z}_T(x_{T|\infty}), \\
&\geq \hat{\mathcal{Z}}_T(x_{T|\infty}).
\end{aligned}$$

Condition **C1** guarantees $\hat{\mathcal{Z}}(x_{T|\infty}) \geq \hat{\phi}_T$ and, therefore, $\sigma(\|x_0 - \hat{x}_0\|) \geq \hat{\phi}_T$. Let us now assume $\mathcal{Z}_{T-N}(x_{T-N|\infty}) \geq \hat{\mathcal{Z}}_{T-N}(x_{T-N|\infty})$ for the induction argument. Utilizing the optimality principle, we have, for all $T > N$,

$$\begin{aligned}
\sigma(\|x_0 - \hat{x}_0\|) &\geq \min_{z, \{w_k\}_{k=T-N}^{T-1}} \Bigg\{ \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \mathcal{Z}_{T-N}(z) \ : \\
&\qquad\qquad (z, \{w_k\}) \in \Omega_T^N,\ x(T; z, T-N, \{w_j\}) = \ x_{T|\infty} \Bigg\}, \\
&= \mathcal{Z}_T(x_{T|\infty}), \ \text{(by optimality)} \\
&\geq \min_{z, \{w_k\}_{k=T-N}^{T-1}} \Bigg\{ \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \hat{\mathcal{Z}}_{T-N}(z) \ : \\
&\qquad\qquad (z, \{w_k\}) \in \Omega_T^N,\ x(T; z, T-N, \{w_j\}) = \ x_{T|\infty} \Bigg\}, \\
&\geq \hat{\mathcal{Z}}_T(x_{T|\infty}). \ \text{(by the induction assumption and } \mathbf{C2}\text{)}
\end{aligned}$$

Condition **C1** guarantees $\hat{\mathcal{Z}}_T(x_{T|\infty}) \geq \hat{\phi}_T$ for all $T \geq 0$. The sequence $\{\hat{\phi}_k\}$, therefore, is monotone nondecreasing and bounded above by $\sigma(\|x_0 - \hat{x}_0\|)$. Hence, it is convergent, and the partial sum

$$\sum_{k=T-N}^{T-1} L_k(\hat{w}^{\text{mh}}_{k|T-1}, \hat{v}^{\text{mh}}_{k|T-1}) \to 0,$$

as $T \to \infty$, because the summation in (3.6) is nonnegative. Lemma 3.3.2 guarantees the estimation error $\|x(T; x_0, 0) - \hat{x}_T^{\text{mh}}\| \to 0$ as claimed.

To prove stability, let $\epsilon > 0$ and choose $\varrho > 0$ as specified by Lemma 3.3.2 such that if

$$\sum_{k=T-N}^{T-1} L_k(\hat{w}_{k|T-1}^{\text{mh}}, \hat{v}_{k|T-1}^{\text{mh}}) \leq \varrho,$$

then $\|x(T; x_0, 0) - \hat{x}_T^{\text{mh}}\| \leq \epsilon$ for all $T \geq N_o$. If we choose $\delta > 0$ such that $\delta < \sigma^{-1}(\varrho)$ (the existence of $\sigma^{-1}(\cdot)$ follows from Fact 3.2.2), then we obtain the following inequality for all $T \geq N \geq N_o$.

$$\begin{aligned}
\sigma(\delta) \quad &\geq \quad \sum_{k=T-N}^{T-1} L_k(\hat{w}_{k|T-1}^{\text{mh}}, \hat{v}_{k|T-1}^{\text{mh}}) + \hat{\mathcal{Z}}_{T-N}(\hat{x}_{T-N|T-1}^{\text{mh}}) \\
&\geq \quad \sum_{k=T-N}^{T-1} L_k(\hat{w}_{k|T-1}^{\text{mh}}, \hat{v}_{k|T-1}^{\text{mh}}).
\end{aligned}$$

Hence, if the initial estimation error $\|x_0 - \hat{x}_0\| \leq \delta$, then the estimation error

$$\|x(T; x_0, 0) - \hat{x}_T^{\text{mh}}\| \leq \epsilon$$

for all $T \geq N$ as claimed. $\square$

When the system dynamics are nonlinear, we are unable in general to construct an approximate arrival cost that satisfies condition **C2** with exception of the obvious choice $\hat{\mathcal{Z}}_T(\cdot) = \hat{\phi}_T$. As the proof of Proposition 3.4.4 demonstrates, condition **C2** is sufficient to guarantee $\sigma(\|x_0 - \hat{x}_0\|)$ is a uniform upper bound to the optimal cost $\hat{\phi}_k$ for all $k \geq 0$. While global satisfaction of the inequality (3.5) in **C2** is ideal, we may circumvent the issue by explicitly ensuring $\sigma(\cdot)$ is a uniform bound in nominal application. Suppose the sequence of approximate arrival costs $\{\hat{\mathcal{Z}}_j(\cdot)\}_{j=0}^{\infty}$ satisfies condition **C1**. The purpose of condition **C2** is to ensure the sequence $\{\hat{\mathcal{Z}}_j(x_{j|\infty})\}$ is monotone nonincreasing (see **A3**):

$$\hat{\mathcal{Z}}_T(x_{T|\infty}) \quad \leq \quad \min_{z, \{w_k\}_{k=T-N}^{T-1}} \Bigg\{ \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \hat{\mathcal{Z}}_{T-N}(z) : \tag{3.7a}$$

$$(z, \{w_k\}) \in \Omega_T^N, \;\; x(T; z, T-N, \{w_j\}) = x_{T|\infty} \Bigg\}, \tag{3.7b}$$

$$\leq \quad \sum_{k=T-N}^{T-1} L_k(w_{k|\infty}, v_{k|\infty}) + \hat{\mathcal{Z}}_{T-N}(x_{T-N|\infty}). \tag{3.7c}$$

Rather than rely on the general structure of the sequence $\{\hat{\mathcal{Z}}_j(\cdot)\}$ to satisfy the inequality (3.7c), we may force the sequence $\{\hat{\mathcal{Z}}_j(x_{j|\infty})\}$ to be monotone nonincreasing explicitly by scaling the approximate arrival costs

$$\hat{\mathcal{Z}}_j(\cdot) \leftarrow \beta_j \left(\hat{\mathcal{Z}}_j(\cdot) - \hat{\phi}_j \right) + \hat{\phi}_j$$

where $\beta_j \in [0, 1]$.

If we knew the sequence $\{x_{k|\infty}\}_{k=0}^{\infty}$ defined in **A3**, then enforcing the inequality (3.7c) is easy. It is sufficient to scale $\hat{\mathcal{Z}}_T(\cdot)$ such that the inequality (3.7c) is satisfied. The problem is that we rarely know of a sequence satisfying **A3** *a priori* without first solving a full information estimation problem. However, to satisfy the inequality (3.7c) at time T, we need only to know the last N elements of the sequence $\{x_{k|\infty}, w_{k|\infty}\}_{k=T-N}^{T-1}$. Even this information is unavailable *a priori*, though we may obtain it online. What we need to generate online is a feasible state sequence $\{x_k^0, w_k^0\}_{k=T-N}^{T-1}$ that is bounded by the initial estimation error in nominal application. We can generate this feasible sequence using $\hat{\mathcal{Z}}_{T-N}(\cdot) = \hat{\phi}_{T-N}$. Recall from Remark 3.4.1 that this choice for the approximate arrival cost yields a stable constrained observer. Once we have a feasible sequence, we can scale $\hat{\mathcal{Z}}_T(\cdot)$ such that it satisfies (3.7c).

Consider the MHE problem where we choose $\hat{\mathcal{Z}}_T(\cdot) = \hat{\phi}_T$. We formulate this estimation problem as the following optimal control problem[4]

$$P_3(T): \qquad \psi_T^* = \min_{z, \{w_k\}_{k=T-N}^{T-1}} \left\{ \sum_{k=T-N}^{N-1} L_k(w_k, v_k) \ : \ (z, \{w_k\}) \in \Omega_T^N \right\}.$$

For $T \leq N$, $P_3(T)$ is defined to be $P_1(T)$. The solution to $P_3(T)$ is the pair

$$\left(z^*, \{\hat{w}_{k|T-1}^0\}_{k=T-N}^{T-1} \right),$$

and that optimal pair yields an estimate $\{\hat{x}_{k|T-1}^0\}_{k=T-N}^{T}$ of of the actual sequence $\{x_k\}$, where

$$\hat{x}_{k|T-1}^0 := x(k, z^*, T-N, \{\hat{w}_{j|T-1}^0\}).$$

It follows that $\hat{x}_0^0 = \hat{x}_0$. We formulate the estimation strategy as the following algorithm.

Estimation algorithm 1

Data $N \in \mathbb{N}$.

Initialization: For $T \leq N$ do:

1. Solve $P_2(T)$ to generate $\{\hat{x}_k\}_{k=1}^N$ and $\{\hat{\phi}_k\}_{k=1}^N$.
2. Solve $P_3(T)$ to obtain $\hat{x}_{0|N-1}^0$ and $\{\psi_k^*\}_{k=1}^N$.
3. For $k = 1, \ldots, N$, set $U_k = \psi_k^* + \Gamma(\hat{x}_{0|N-1}^0)$.

Step 1 For $T > N$ do:

1. Solve $P_3(T)$ to obtain $\hat{x}_{T-N|T-1}^0$ and ψ_T^*.
2. Set $U_T = \psi_T^* + U_{T-N}$.
3. Construct $\hat{\mathcal{Z}}_{T-N}(\cdot)$ so that it satisfies **C1**.
4. Set

$$\beta_{T-N} = \max_{\beta \in [0,1]} \left\{ \beta \ : \ \beta \left(\hat{\mathcal{Z}}_{T-N}(\hat{x}_{T-N|T-1}^0) - \hat{\phi}_{T-N} \right) + \hat{\phi}_{T-N} \leq U_{T-N} \right\}.$$

5. Set

$$\hat{\mathcal{Z}}_{T-N}(\cdot) \leftarrow \beta_{T-N} \left(\hat{\mathcal{Z}}_{T-N}(\cdot) - \hat{\phi}_{T-N} \right) + \hat{\phi}_{T-N}$$

6. Solve $P_2(T)$ and obtain $\hat{x}_T$ and $\hat{\phi}_T$.

[4] Adding a constant to the objective function does not affect the answer. For simplicity, we choose $\hat{\mathcal{Z}}_T(\cdot) = 0$.

Step 2 Let $T \leftarrow T+1$. Go to Step 1.

Remark 3.4.5 *If we choose*

$$\hat{\mathcal{Z}}_j(x) = (x - \hat{x}_j)^T \Pi^{-1}(x - \hat{x}_j) + \hat{\phi}_j,$$

where the matrix Π *is symmetric positive semi-definite, then* **C1** *is automatically satisfied; let* $\bar{\gamma}(\cdot) = (1 + \lambda_{\max}(\Pi^{-1}))\|(\cdot)\|^2$.

To prove stability, we require the following lemma.

Lemma 3.4.6 *Let* $x_0 \in \mathbb{X}_0$. *Suppose* **A0**–**A3** *are true, the system (3.1) is uniformly observable, and* $N \geq N_o$*, then there exists K-functions* $\theta_1(\cdot)$ *and* $\theta_2(\cdot)$ *such that*

$$\|x(T; x_0, 0) - \hat{x}_T\| \leq \theta_1 \left(\sum_{k=T-N}^{T-1} L_k(\hat{w}_{k|T-1}, \hat{v}_{k|T-1}) \right)$$

and

$$\|x(T-N; x_0, 0) - \hat{x}_{T-N|T-1}\| \leq \theta_2 \left(\sum_{k=T-N}^{T-1} L_k(\hat{w}_{k|T-1}, \hat{v}_{k|T-1}) \right)$$

for all $T \geq N$.

Proof. The proof is given in Appendix 3.8.3. $\square$

The stability of the proposed algorithm relies on the stability of the estimator defined by $P_3(T)$. We know from Proposition 3.4.4 that $\|\hat{x}_T^0 - x(T; x_0, 0)\| \to 0$ as $T \to \infty$. More importantly, we know also that $\{\psi_k^*\} \in l_2(\mathbb{R})$. This implies the sequence $\{U_k\}$ is bounded.

Proposition 3.4.7 *If assumptions* **A0**–**A3** *hold, the system (3.1) is uniformly observable, and* $N \geq N_o$*, then, for all* $\hat{x}_0 \in \mathbb{X}_0$*, MHE using estimation algorithm 1 is an asymptotically stable observer for the system (3.4).*

Proof. From the preceding arguments (see the proof of Proposition 3.4.4), it suffices to show U_T is bounded uniformly for all $k \geq 0$ by $\|x_0 - \hat{x}_0\|$. Let $V = \sigma(\|x_0 - \hat{x}_0\|) + \Gamma(\hat{x}^0_{0|N-1})$. Optimality guarantees $\sum_{k=T-N}^{T-1} L_k(w_{k|\infty}, v_{k|\infty}) \geq \psi_T^*$ for all $T \geq N$. Hence, by **A3**, we have $U_k \leq V$ for all $k \geq N$. By construction, for $T \geq N$, $\hat{\mathcal{Z}}_{T-N}(\hat{x}^0_{T-N|T-1}) \leq U_{T-N}$. Because $\left(\hat{x}^0_{T-N|T-1}, \{\hat{w}^0_{k|T-1}\}\right) \in \Omega_T^N$, optimality implies $\hat{\phi}_T \leq U_T$. Hence, the sequence $\{\hat{\phi}_j\}$ is bounded above by V and, consequently, $\|x(T; x_0, 0) - \hat{x}_T\| \to 0$ as $T \to \infty$.

We now establish that V is bounded by $\|x_0 - \hat{x}_0\|$. By assumption **A3**, $\psi_N^* \leq \sigma(\|x_0 - \hat{x}_0\|)$ and, by Lemma 3.4.6,

$$\begin{aligned} \|\hat{x}^0_{0|N-1} - x_0\| &\leq \theta_2(\psi_N^*), \\ &\leq \theta_2\left(\sigma(\|x_0 - \hat{x}_0\|)\right). \end{aligned}$$

Hence, we obtain

$$\begin{aligned} V &\leq \sigma(\|x_0 - \hat{x}_0\|) + \bar{\gamma}(\|\hat{x}^0_{0|N-1} - \hat{x}_0\|), \\ &\leq \sigma(\|x_0 - \hat{x}_0\|) + \bar{\gamma}\left(\|\hat{x}^0_{0|N-1} - x_0\| + \|x_0 - \hat{x}_0\|\right), \\ &\leq \sigma(\|x_0 - \hat{x}_0\|) + \bar{\gamma}\left(\theta_2\left(\sigma(\|x_0 - \hat{x}_0\|)\right) + \|x_0 - \hat{x}_0\|\right), \\ &:= \omega(\|x_0 - \hat{x}_0\|), \end{aligned}$$

where $\bar{\gamma}(\cdot)$ results from applying condition **C1** and $\omega(\cdot)$ is a K-function. The existence of the K-function $\omega(\cdot)$ follows from Fact 3.2.3. $\square$

Remark 3.4.8 *Estimation algorithm 1 is always solvable even in the presence of noise. Optimality implies $\hat{\phi}_T \leq U_T$. Therefore, any $\hat{\mathcal{Z}}_j^0(\cdot)$ satisfying condition* **C1** *guarantees the existence of a $\beta_j \geq 0$.*

Example 3.4.9 (A Simple Example Demonstrating Estimation Strategy 1) *Consider the scalar system*

$$x_{k+1} = 1.1x_k + w_k, \; y_k = x_k + v_k.$$

This system has been studied by Muske and Rawlings (1995) and Findeisen (1997) as an example where MHE is unstable. The following least squares objective function

$$\hat{\phi}_T = \sum_{k=T-N}^{T-1} w_k^2 + \frac{v_k^2}{100} + \frac{(x_{T-N} - \hat{x}_{T-N})^2}{P}$$

is used, where the initial penalty $\hat{\mathcal{Z}}_j(x_j) = (x_j - \hat{x}_j)^2/P$. We assume the initial condition is $x_0 = 0$ for the process and $\hat{x}_0 = 1$ for the model. For a horizon of $N = 5$, $P \geq 4.3$ was necessary for stability. For a horizon of $N = 10$, $P \geq 1.05$ was necessary for stability.

A comparison of the MHE with and without the proposed scaling strategy is shown in Figures 3.2 and 3.3. Without the scaling of the initial penalty, the MHE is unstable. The response of the Kalman filter or full information solution is plotted as a benchmark. As evident from the figures, the scaling parameter β is reduced until the MHE is stable. Figure 3.4 shows a comparison of MHE with and without the scaling strategy when the unscaled initial penalty is sufficiently small to guarantee stability. In Figure 3.4, the scaling parameter $\beta = 1$.

The cycling effect evident in all of the figures is due to the filter update used: the dynamics of the estimator at times $T - N-$ to $T - 1$ are unrelated to the dynamics of the estimator at times $T - N$ and T. One strategy to remove the cycling effect is to employ a smoothing update. Findeisen (1997) provides a comprehensive discussion of the cycling effect. Smoothing strategies are discussed in Chapter 4

We desire $\beta_T = 1$ when $\hat{\mathcal{Z}}_T(\cdot)$ satisfies condition **C2**. If we assume $x(k; x_0, 0) \in \mathbb{X}_k$ for all $k \geq 0$, then optimality and the observability assumption imply $\hat{x}^0_{T-N|T-1} = x(T - N; x_0, 0)$ for all $T \geq N$ and, as a consequence, $U_T = 2\Gamma(x_0)$. It follows by optimality and condition **C2** that for $T \geq N$

$$\begin{aligned} \hat{\mathcal{Z}}_T(\hat{x}^0_{T-N|T-1}) &= \hat{\mathcal{Z}}_T(x(T-N; x_0, 0)), \\ &\leq 2\Gamma(x_0). \end{aligned}$$

Therefore, $\beta_T = 1$. When the constraints only satisfy **A3** or when we consider suboptimal algorithms, estimation algorithm 1 does not guarantee $\beta_T = 1$ when the sequence $\{\hat{\mathcal{Z}}^0_T(\cdot)\}$ satisfies condition **C2**. To guarantee $\beta_T = 1$, we need to modify estimation algorithm 1. We need specifically to construct a state and disturbance sequence $\{x_k, w_k\}_{k=0}^{\infty}$ satisfying the state equation (3.1) and constraints. The natural choice for the state sequence is $\{\hat{x}^0_k\}$. However, there may not exist a complementary disturbance sequence $\{\hat{w}^0_k\}$ satisfying the state equation (3.1) and constraints (i.e. $\left(\hat{x}^0_0, \{\hat{w}^0_k\}\right) \in \Omega_T$ for all $T \geq 0$). Therefore, we need to make the following assumption.

A4 For all $x \in \mathbb{X}_k$, $z \in \mathbb{X}_{k+1}$, and $k \geq 0$, there exists some $w \in \mathbb{W}_k$ and a K-function $\kappa(\cdot)$ such that

1. $z = f_k(x, w)$,

2. $\kappa(\|f(x, w) - f_k(x, 0)\|) \geq \|w\|$.

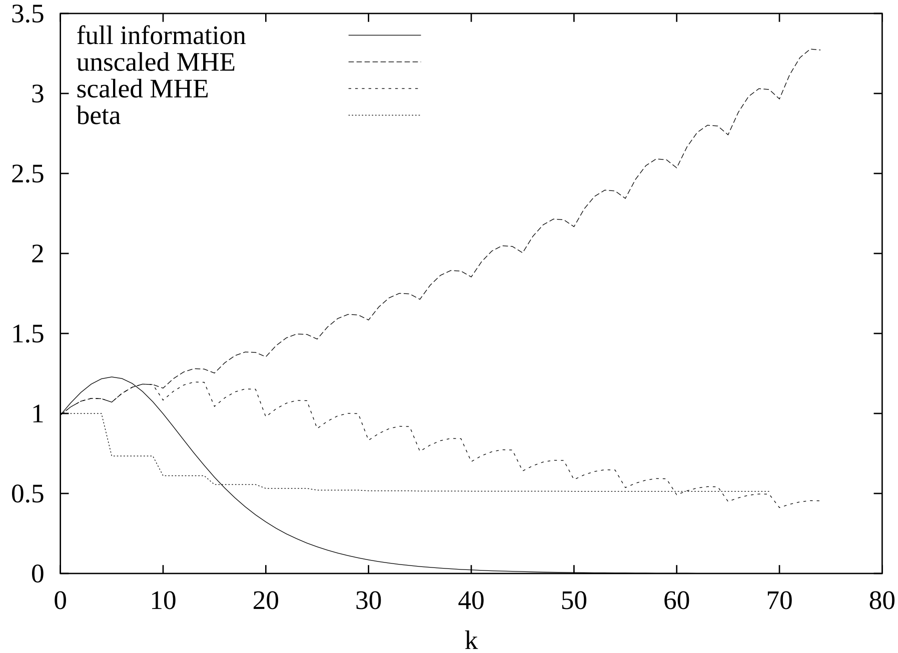


Figure 3.2: Estimation error for $M = 5$ and $P = 3$

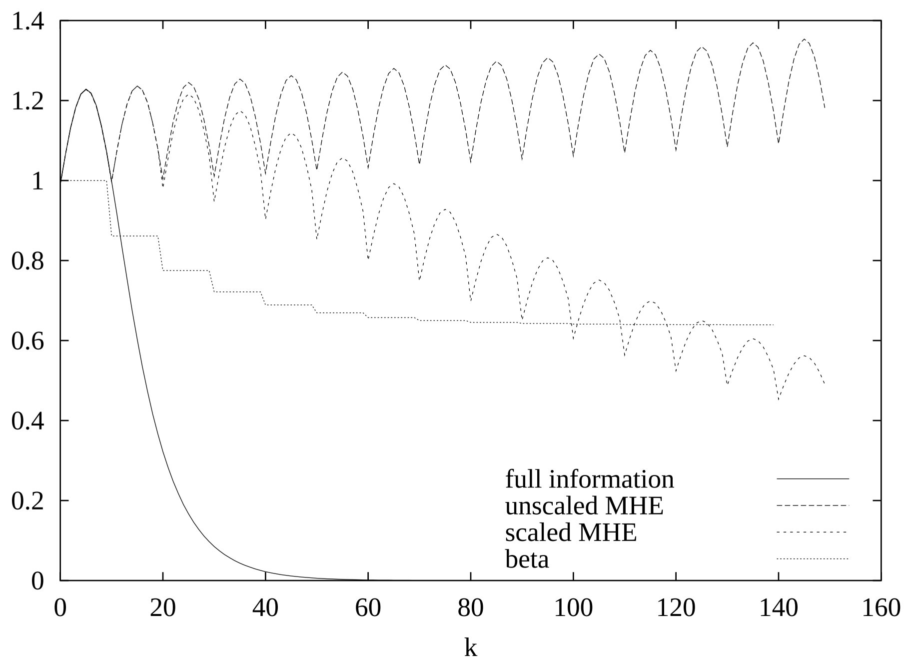


Figure 3.3: Estimation error for $M = 10$ and $P = 1$

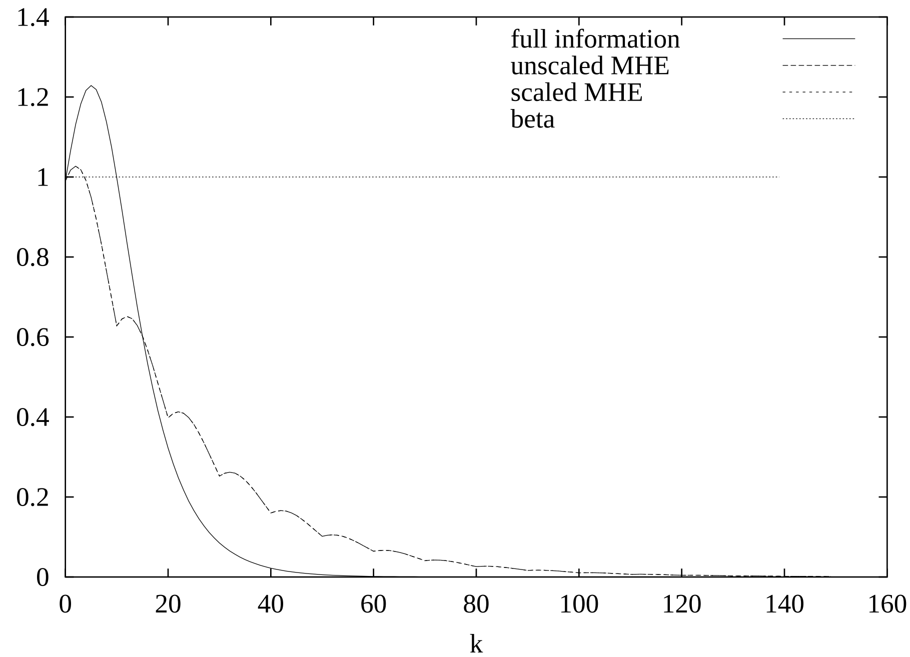


Figure 3.4: Estimation error for $M = 10$ and $P = 5$

Assumption **A4** is easy to verify though restrictive in our framework; **A4** typically requires $f_k(x, w) = f_k(x) + w$. We hope, ideally, that the closed-loop state and disturbance estimates generate a feasible trajectory, in that the transition from $\hat{x}_k^0$ to $\hat{x}_{k+1}^0$ is feasible. However, nowhere in our proposed framework is this assumption necessary or expected. The full information and moving horizon estimator implicitly recalculate the entire state and disturbance trajectory at each time interval, so, in either formulation, feasibility of the closed-loop estimates is not directly addressed.

Estimation algorithm 2

Data $N \in \mathbb{N}$.

Initialization: For $T \leq N$;

1. Solve $P_2(T)$ to generate $\{\hat{x}_T\}_{k=1}^T$ and $\{\hat{\phi}_T\}_{k=1}^T$.
2. Solve $P_3(T)$ to obtain $\hat{x}_{0|T-1}^0$, $\{\hat{x}_k^0\}_{k=1}^T$, and $\{\psi_k^*\}_{k=1}^N$.
3. Set $U_k = \Gamma(\hat{x}_{0|N-1}^0) + \psi_N^*$ for $k = 1, \ldots, N$.

Step 1 For $T > N$ do:

1. Solve $P_3(T)$ to obtain $\hat{x}_T^0$.
2. Choose $\bar{w}_{T-1}$ such that

$$\hat{x}_T^0 = f_{T-1}(\hat{x}_{T-1}^0, \bar{w}_{T-1}).$$

3. Set

$$U_T = L_{T-1}(\bar{w}_{T-1}, \hat{v}^0_{T-1}) + U_{T-1},$$

where $\hat{v}^0_{T-1} = y_{T-1} - h_{T-1}(\hat{x}^0_{T-1})$.

4. Construct $\hat{\mathcal{Z}}_{T-N}(\cdot)$ so that it satisfies **C1**.

5. Solve the problem

$$\beta_{T-N} = \max_{\beta\in[0,1]} \left\{\beta \ : \ \beta\left(\hat{\mathcal{Z}}_{T-N}(\hat{x}^0_{T-N}) - \hat{\phi}_{T-N}\right) + \hat{\phi}_{T-N} \leq U_{T-N}\right\}.$$

6. Let

$$\hat{\mathcal{Z}}_{T-N}(\cdot) \leftarrow \beta_{T-N}\left(\hat{\mathcal{Z}}_{T-N}(\cdot) - \hat{\phi}_{T-N}\right) + \hat{\phi}_{T-N}$$

7. Solve $P_2(T)$ and obtain $\hat{x}_T$ and $\hat{\phi}_T$.

Step 2 Let $T \leftarrow T+1$. Go to Step 1.

Lemma 3.4.10 *If assumptions* **A0–A4** *hold, the system (3.4) is uniformly observable, and* $N \geq N_o$*, then there exists a K-function* $\mu(\cdot)$ *such that*

$$L_k(\bar{w}_k, \hat{v}^0_k) \leq \mu(\psi^*_k) \vee \mu(\psi^*_{k+1})$$

for all $k \geq N$*.*

Proof. The proof is given in Appendix 3.8.4. □

Proposition 3.4.11 *If assumptions* **A0–A4** *hold, the K-function* $\mu(\cdot)$ *defined in Lemma 3.4.10 is locally Lipschitz continuous at the origin, the system (3.1) is uniformly observable, and* $N \geq N_o$*, then, for all* $\hat{x}_0 \in \mathbb{X}_0$*, MHE using the estimation algorithm 2 is an asymptotically stable observer for the system (3.4).*

Proof. From the preceding arguments (see the proofs of Propositions 3.4.4 and 3.4.7), it suffices to show the quantity

$$U_\infty = \lim_{T\to\infty} \sum_{k=N}^{T-1} L_k(\bar{w}_k, \hat{v}^0_k)$$

is bounded by $\|x_0 - \hat{x}_0\|$. By assumption, there exists an $\epsilon > 0$ and $K > 0$ such that for all $x \in B_\epsilon, \mu(x) \leq K\|x\|$. Optimality and assumption **A3** guarantee $\sum_{k=N}^{\infty} \psi^*_k \leq N\sigma(\|x_0 - \hat{x}_0\|)$. Therefore, there exists $M > 0$ such that $\sum_{k=M}^{\infty} \psi^*_k \leq \epsilon$ and, consequently, $\sum_{k=M}^{\infty} L_k(\bar{w}_k, \hat{v}^0_k) \leq 2NK\sigma(\|x_0 - \hat{x}_0\|)$. The summation $\sum_{k=N}^{\infty} L_k(\bar{w}_k, \hat{v}^0_k)$ is convergent and, hence,

$$\sum_{k=N}^{\infty} L_k(\bar{w}_k, \hat{v}^0_k) \leq 2\mu(N\sigma(\|x_0 - \hat{x}_0\|)).$$

Therefore, the proposition follows as claimed. □

Remark 3.4.12 *If we assume the K-functions $\gamma(\cdot)$ defined in* **A1** *and $\kappa(\cdot)$ defined in* **A4** *are locally Lipschitz continuous at the origin, the stage cost functions are quadratic (i.e. $L_k(w, v) = w^T Q_k w + v_k^T R v_k$), and $\varphi(\cdot) = \alpha \|(\cdot)\|^2$ for some $\alpha > 0$ where the K-function $\varphi(\cdot)$ is defined in Definition 3.2.4, then the K-function $\mu(\cdot)$ defined in Lemma 3.4.10 is locally Lipschitz continuous at the origin.*

Proposition 3.4.13 *Suppose* **A0–A4** *are true, the sequence $\{\hat{\mathcal{Z}}_j(\cdot)\}$ satisfies* **C2***, the system (3.1) is uniformly observable, and $N \geq N_o$. Then, for all $x_0 \in \mathbb{X}_0$, $\beta_T = 1$ in estimation algorithm 2 for all $T \geq 0$.*

Proof. We proceed using an induction argument. For $T \leq N$, we have

$$\hat{\mathcal{Z}}_T^0(\hat{x}_T^0) \leq \min_{x_0, \{w_k\}_{k=0}^{T-1}} \left\{ \Phi_T(x_0, \{w_k\}) \ : \ \begin{array}{c} (x_0, \{w_k\}) \in \Omega_T \\ x(T; x_0, 0, \{w_j\}) = \hat{x}_T^0 \end{array} \right\} \leq U_T.$$

Now consider $T > N$ and assume $\hat{\mathcal{Z}}_{T-N}^0(\hat{x}_{T-N}^0) \leq U_{T-N}$ for the induction argument. Then, we obtain the following inequalities

$$\begin{aligned} U_T \ &\geq \ \sum_{k=N}^{T-1} L_k(\hat{v}_k^0, \bar{w}_k) + \hat{\mathcal{Z}}_{T-N}(\hat{x}_{T-N}^0), \\ &\geq \ \min_{\{w_k\}_{k=T-N}^{T-1}} \left\{ \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \hat{\mathcal{Z}}_{T-N}(z) \ : \right. \\ &\qquad \left. (z, \{w_k\}) \in \Omega_T^N, \ \begin{array}{c} z = \hat{x}_{T-N}^0 \\ x(T; z, T-N, \{w_k\}) = \hat{x}_T^0 \end{array} \right\}, \\ &\geq \ \min_{z, \{w_k\}_{k=T-N}^{T-1}} \left\{ \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \hat{\mathcal{Z}}_{T-N}(z) \ : \right. \\ &\qquad \left. (z, \{w_k\}) \in \Omega_T^N, \ x(T; z, T-N, \{w_k\}) = \hat{x}_T^0 \right\}, \\ &\geq \ \hat{\mathcal{Z}}_T(\hat{x}_T^0). \end{aligned}$$

Hence, the proposition follows. □

The strength of estimation strategy 2 over estimation strategy 1 is clearly the result of Proposition 3.4.13: if $\{\hat{\mathcal{Z}}_j^0(\cdot)\}$ satisfies assumption **C2**, then $\beta_T = 1$. The weakness is clearly assumption **A4** and the assumption of local Lipschitz continuity at the origin.

Remark 3.4.14 *The system (3.4) need not generate the data for the result of Proposition 3.4.13 to hold. The result holds regardless of which system generates the data. Furthermore, estimation algorithm 2 is always solvable even in the presence of noise. See Remark 3.4.8.*

3.5 Suboptimal Implementation

To guarantee stability in the proposed state estimation strategies, we require a global solution to both $P_2(T)$ and $P_3(T)$. This computational requirement presents a formidable barrier to online implementation. Aside from the computational burden, optimization may not yield global solutions unless the problem is convex. Convexity is restrictive, because the functions $f_k(\cdot)$ and $h_k(\cdot)$ describing the system (3.1) need to be affine. Without convexity, we can reasonably expect only a *local* solution to $P_2(T)$ and $P_3(T)$. Strategies exist for finding a global solution, though they are currently impractical for online implementation. The difficulty in global optimization is not finding a solution, but rather verifying whether a particular solution is global. Unless global information such as lower bounds or Lipschitz constants are available, one needs to sample a dense subset of the decision space in order to guarantee a particular solution is global (Stephens and Baritompa 1998). As global information is rarely available, we need to adjust the MHE algorithm to allow for local solutions. However, even finding a local solution may be impractical for online implementation; often the optimization algorithm may not converge in the requisite time. We would prefer, instead of requiring a local solution, if a cost reduction is sufficient to guarantee stability. We refer to a solution that yields only a feasible cost reduction as a suboptimal solution. Mayne (1995*b*) provides a discussion of the computational issues in receding horizon control (RHC), and most of these issues translate directly to MHE.

One potential solution for reducing the computational complexity of MHE is the suboptimal strategy for RHC first proposed by Michalska and Mayne (1993) and further developed in discrete time by Scokaert, Mayne and Rawlings (1999). Instead of requiring a global minimum to the optimization problem in RHC, the authors demonstrate a feasible trajectory is sufficient for stability. Further improvements (i.e. a reduced cost) in the control are possible during the sampling period if time permits, but they are not necessary for stability. The idea substantially reduces the complexity of the control problem, because closed-loop stability does not necessitate a global or even a local solution to the optimization problem. This result is also applicable to MHE. A solution to $P_3(T)$ establishes feasibility in the state estimation problem. So long as U_T bounds suboptimal value $\hat{\phi}'_T$ of the cost function in $P_2(T)$, the estimator is stable. Between the sampling periods, the optimizer may attempt to improve the value of $\hat{\phi}'_T$. However, a local or global minimum is not required. While this reduces the computational burden involved in solving $P_2(T)$, we still require a global solution to $P_3(T)$.

One strategy to relax the requirement of a global solution to $P_3(T)$ is the moving horizon observer proposed by Michalska and Mayne (1995). Instead of requiring a global solution, they propose to reduce exponentially the cost $\sum_{k=T-N}^{T-1} L_k(w_k, v_k)$ by imposing the reduction condition $\psi'_T \leq \rho\psi'_{T-N}$ for some scalar $\rho \in (0,1)$, where ψ'_T denotes the suboptimal value of the cost function. Without constraints it is straightforward to translate their result to MHE, because $P_3(T)$ is equivalent to the Mayne and Michalska observer. We could impose a related reduction condition on the cost function in $P_3(T)$ for MHE, and then scale the terminal penalty appropriately. This gradual reduction allows us to circumvent the need for a global minimum to $P_3(T)$. When constraints are present, it is not always possible to reduce the cost function in $P_3(T)$, even in nominal application. Otherwise, when constraints are not present, $\psi^*_T = 0$ and we can always satisfy the condition $\psi'_T \leq \rho\psi'_{T-N}$. One can account for the limitations imposed by constraints in nominal application by adding a sequence of decreasing slack variables.

Before presenting the two suboptimal strategies, we provide first our definition of a suboptimal solution.

Definition 3.5.1 *For $T \geq N$, we say the pair $\left(\hat{x}^{mh}_{T-N|T-1}, \{\hat{w}^{mh}_k\}_{k=T-N|T-1}^{T-1}\right)$ is a* **suboptimal** *solution to $P_2(T)$ with respect to some property* **S** *if the pair is an element of Ω_T^N and satisfies the property* **S**. *Furthermore, let*

$$\hat{\phi}'_T = \sum_{T-N}^{T-1} L(\hat{w}^{mh}_{k|T-1}, \hat{v}^{mh}_{k|T-1}) + \hat{\mathcal{Z}}_{T-N}(\hat{x}^{mh}_{T-N|T-1})$$

denote the suboptimal cost.

Definition 3.5.2 *For $T \geq N$, we say the pair $\left(\hat{x}^{0}_{T-N|T-1}, \{\hat{w}^{0}_k\}_{k=T-N|T-1}^{T-1}\right)$ is a* **suboptimal** *solution to $P_3(T)$ with respect to some property* **S** *if the pair is an element of Ω_T^N and satisfies the property* **S**.

Furthermore, let

$$\psi'_T = \sum_{T-N}^{T-1} L(\hat{w}^0_{k|T-1}, \hat{v}^0_{k|T-1})$$

denote the suboptimal cost.

Suboptimal estimation algorithm 1

Data $N \in \mathbb{N}$, $\rho \in (0,1)$, and $\{s_k\}_{k=0}^{\infty} \in l_2(\mathbb{R}_{\geq 0})$

Initialization: For $T \leq N$ do:

1. Solve (suboptimally) $P_2(T)$ to generate $\{\hat{x}_k\}_{k=1}^N$ and $\{\hat{\phi}'_k\}_{k=1}^N$.
2. Solve (suboptimally) $P_3(T)$ to obtain $\hat{x}^0_{0|N-1}$ and $\{\psi'_k\}_{k=1}^N$.
3. For $k = 1, \ldots, N$, set $U_k = \psi'_k + \Gamma(\hat{x}^0_{0|N-1})$.

Step 1 For $T > N$ do:

1. Solve $P_3(T)$ suboptimally with respect to the inequality constraint
$$\psi'_T \leq \rho \psi'_{T-N} + s_T,$$
to obtain $\hat{x}^0_{T-N|T-1}$ and ψ'_T.
2. Set $U_T = \psi'_T + U_{T-N}$.
3. Construct $\hat{\mathcal{Z}}_{T-N}(\cdot)$ so that it satisfies **C1**.
4. Set
$$\beta_{T-N} = \max_{\beta \in [0,1]} \left\{ \beta \ : \ \beta \left(\hat{\mathcal{Z}}_{T-N}(\hat{x}^0_{T-N|T-1}) - \hat{\phi}'_{T-N} \right) + \hat{\phi}'_{T-N} \leq U_{T-N} \right\}.$$
5. Set
$$\hat{\mathcal{Z}}_{T-N}(\cdot) \leftarrow \beta_{T-N} \left(\hat{\mathcal{Z}}_{T-N}(\cdot) - \hat{\phi}'_{T-N} \right) + \hat{\phi}'_{T-N}$$
6. Solve $P_2(T)$ suboptimally with respect to the constraint
$$\hat{\phi}'_T \leq U_T$$
to obtain $\hat{x}_T$ and $\hat{\phi}'_T$.

Step 2 Let $T \leftarrow T + 1$. Go to Step 1.

The choice of the slack sequence $\{s_k\}$ is the challenge in implementing either one of the suboptimal strategies. One possibility is an exponentially decreasing sequence. For example, $s_k = s^k$ for some $s \in (0,1)$. The following proposition proves the existence of a bounded slack sequence.

Proposition 3.5.3 *If assumptions* **A0–A3** *hold, the system (3.1) is uniformly observable, and* $N \geq N_o$*, then there exists a sequence* $\{s_k\}_{k=0}^{\infty} \in l_2(\mathbb{R}_{\geq 0})$ *and a suboptimal solution to* $P_3(T)$ *such that*

$$\psi'_T \leq \rho\psi'_{T-N} + s_T.$$

Proof. We know from Proposition 3.4.4 that the series $\sum_{k=0}^{\infty} \psi_k^* < \infty$. If we choose $s_k = \psi_k^*$, then the proposition follows by inspection. □

In order to guarantee stability for suboptimal estimation strategy 1, we require the following assumption that states if the initial estimation error is small, then the suboptimal solution is also small.

A5 Suppose the system (3.4) with initial condition x_0 generates the data (i.e. $y_k = y(k; x_0, 0)$). There exists a K-function $\sigma_1(\cdot)$ such that, for $N \geq N_o$,

$$\sum_{k=0}^{N-1} L_k(\hat{w}_{k|N}^0, \hat{v}_{k|N}^0) \leq \sigma_1(\|x_0 - \hat{x}_0\|)$$

where the sequence $\{\hat{w}_{k|N}^0, \hat{v}_{k|N}^0\}_{k=0}^{N-1}$ denote the suboptimal solution to $P_3(N)$. Furthermore, there exists a K-function $\sigma_2(\cdot)$ and a sequence $\{s_k\} \in l_2(\mathbb{R}_{\geq 0})$ such that $\psi'_T \leq \rho\psi'_{T-N} + s_T$ for all $T \geq N$ and

$$\sum_{k=0}^{\infty} s_k \leq \sigma_2(\|x_0 - \hat{x}_0\|).$$

Proposition 3.5.4 *If assumptions* **A0–A3** *and* **A5** *hold, the system (3.1) is uniformly observable, and* $N \geq N_o$*, then MHE using suboptimal algorithm 1 is an asymptotically stable observer for the system (3.4).*

Proof. From the preceding arguments (see the proofs of Proposition 3.4.4 and 3.4.7), it suffices to show U_T is bounded uniformly by $\|x_0 - \hat{x}_0\|$ for all $T \geq 0$. For arbitrary j, we have the following inequality

$$\psi'_{jN} \leq \rho^{j-1}\psi'_N + \sum_{k=2}^{j} \rho^{j-k} s_{kN}.$$

This inequality implies that the series

$$\sum_{j=1}^{\infty} \psi'_{jN} \leq \sum_{j=0}^{\infty} \rho^j \psi'_N + \sum_{j=0}^{\infty}\sum_{k=2}^{j} \rho^k s_{lN} \tag{3.8a}$$

$$\leq \frac{1}{1-\rho}\left(\psi'_N + \sum_{k=2}^{\infty} s_{kN}\right) < \infty. \tag{3.8b}$$

Using assumption **A5** and (3.8b) we have

$$\sum_{j=1}^{\infty} \psi'_{jN} \leq \frac{1}{1-\rho}\left(\sigma_1(\|x_0 - \hat{x}_0\|) + \sigma_2(\|x_0 - \hat{x}_0\|)\right),$$

and the proposition follows as claimed. □

Suboptimal estimation algorithm 2

Data $N \in \mathbb{N}$, $\rho \in (0,1)$, and $\{s_k\}_{k=0}^{\infty} \in l_2(\mathbb{R}_{\geq 0})$.

Initialization: For $T \leq N$;

1. Solve (suboptimally) $P_2(T)$ to generate $\{\hat{x}_T\}_{k=1}^{T}$ and $\{\hat{\phi}_T\}_{k=1}^{T}$.
2. Solve (suboptimally) $P_3(T)$ to obtain $\hat{x}^0_{0|T-1}$, $\{\hat{x}^0_k\}_{k=1}^{T}$, and $\{\psi'_k\}_{k=1}^{N}$.
3. Set $U_k = \Gamma(\hat{x}^0_{0|N-1}) + \psi'_N$ for $k = 1, \ldots, N$.

Step 1 For $T > N$ do:

1. Solve $P_3(T)$ suboptimally with respect to the inequality constraint
$$\psi'_T \leq \rho\psi'_{T-N} + s_T,$$
to obtain $\hat{x}^0_T$ and ψ'_T.
2. Choose $\bar{w}_{T-1}$ such that
$$\hat{x}^0_T = f_{T-1}(\hat{x}^0_{T-1}, \bar{w}_{T-1}).$$
3. Set
$$U_T = L_{T-1}(\bar{w}_{T-1}, \hat{v}^0_{T-1}) + U_{T-1},$$
where $\hat{v}^0_{T-1} = y_{T-1} - h_{T-1}(\hat{x}^0_{T-1})$.
4. Construct $\hat{\mathcal{Z}}_{T-N}(\cdot)$ so that it satisfies **C1**.
5. Solve the problem
$$\beta_{T-N} = \max_{\beta \in [0,1]} \left\{\beta \ : \ \beta\left(\hat{\mathcal{Z}}_{T-N}(\hat{x}^0_{T-N}) - \hat{\phi}'_{T-N}\right) + \hat{\phi}'_{T-N} \leq U_{T-N}\right\}.$$
6. Let
$$\hat{\mathcal{Z}}_{T-N}(\cdot) \leftarrow \beta_{T-N}\left(\hat{\mathcal{Z}}_{T-N}(\cdot) - \hat{\phi}'_{T-N}\right) + \hat{\phi}'_{T-N}$$
7. Solve $P_2(T)$ suboptimally with respect to the constraint
$$\hat{\phi}'_T \leq U_T$$
to obtain $\hat{x}_T$ and $\hat{\phi}'_T$.

Step 2 Let $T \leftarrow T + 1$. Go to Step 1.

Lemma 3.5.5 *If assumptions* **A0**–**A4** *hold, the system (3.4) is uniformly observable, and* $N \geq N_o$, *then there exists a K-function* $\mu(\cdot)$ *such that*
$$L_k(\bar{w}_k, \hat{v}^0_k) \leq \mu(\psi'_k) \vee \mu(\psi'_{k+1})$$
for all $k \geq N$.

Proof. The proof is identical to Lemma 3.4.10. □

Proposition 3.5.6 *If assumptions* **A0–A5** *hold, the K-function $\mu(\cdot)$ defined in Lemma 3.5.5 is locally Lipschitz continuous at the origin, the system (3.1) is uniformly observable, and $N \geq N_o$, then MHE using suboptimal strategy 2 is an asymptotically stable observer for the system (3.4).*

Proof. It suffices in light of Propositions 3.4.11 to establish

$$U_\infty = \lim_{T\to\infty} \sum_{k=N}^{T-1} L_k(\bar{w}_k, \hat{v}_k^0)$$

is bounded by $\|x_0 - \hat{x}_0\|$. By assumption (see Proposition 3.5.4)

$$\sum_{k=N}^{\infty} \psi'_k \leq \frac{N}{1-\rho}\left(\sigma_1(\|x_0 - \hat{x}_0\|) + \sigma_2(\|x_0 - \hat{x}_0\|)\right).$$

From Lemma 3.5.5, we have

$$\begin{aligned} \sum_{k=N}^{\infty} L_k(\bar{w}_k, \hat{v}_k^0) &\leq 2\mu\left(\frac{N}{1-\rho}\left(\sigma_1(\|x_0 - \hat{x}_0\|) + \sigma_2(\|x_0 - \hat{x}_0\|)\right)\right), \\ &:= c(\|x_0 - \hat{x}_0\|), \end{aligned}$$

and the lemma follows as claimed. □

Proposition 3.5.7 *If assumptions* **A0–A5** *hold, the K-function $\mu(\cdot)$ defined in Lemma 3.5.5 is locally Lipschitz continuous at the origin, the sequence $\{\hat{\mathcal{Z}}_j(\cdot)\}$ satisfies* **C2**, *the system (3.1) is uniformly observable, and $N \geq N_o$, then $\beta_T = 1$ in suboptimal estimation algorithm 2 for all $T \geq 0$.*

For real systems where measurement and state noise are present, we cannot implement either suboptimal strategies 1 or 2; it is unreasonable to assume a decreasing cost. However, the goal of the preceding discussion was *not* to generate an implementable strategy. Rather, our motive was to demonstrate the possibility of deriving sufficient conditions for a stable suboptimal estimation strategy. In application one should solve problems $P_2(T)$ and $P_3(T)$ to the best of one's resources given the particular algorithmic and time constraints. If optimization yields satisfactory solutions, then we may expect reasonable performance in light of Propositions 3.5.4 and 3.5.6. In our experience, optimizers yield excellent results even without any guarantees of global optimality.

The difficulty in implementing the suboptimal strategies is our definition of stability. We cannot realize asymptotic stability in application. Due to state noise, we do not expect the state estimate to converge to the true state of the system even with global solutions to $P_2(T)$ and $P_3(T)$. Our goal in modeling state noise is to account for this fact. A possible question at this point is why have we concentrated our efforts on demonstrating stability under different conditions. The answer is simple; we are concerned with the lack of stability. While stability is not synonymous with performance, the stability guarantees provide a degree of confidence that the algorithm is not structurally flawed. The practitioner may feel uneasy implementing an algorithm with no nominal stability guarantees or one that requires a globally optimal solution to guarantee stability. Furthermore, in the process of deriving sufficient conditions for asymptotic stability under the assumptions of optimality and suboptimality, we identified some key issues in estimator stability. In particular, the estimator should track the data sufficiently and not emphasize the prior information excessively. In order to tackle this problem, we solve $P_3(T)$ and scale the approximate arrival cost $\hat{\mathcal{Z}}_T(\cdot)$ with β_T. We can restate the cost reduction conditions in suboptimal strategies 1 and 2 as requiring our solution of $P_3(T)$ track the data sufficiently and not allow the estimate to move away from the data.

3.6 Receding Horizon Control and Duality

Given the strong duality between the linear quadratic regulator and the Kalman filter, a natural question is whether MHE is the dual to receding horizon control (RHC). We begin by briefly reviewing RHC. See Mayne, Rawlings, Rao and Scokaert (2000) for a more detailed review of constrained RHC.

Suppose the system is modeled by the nonlinear difference equation

$$x_{k+1} = f_k^c(x_k, u_k), \tag{3.9}$$

where $f_k^c : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $f_k^c(0,0) = 0$ for all $k \geq 0$. Let $x(k; z, l, \{u_j\})$ denote the solution of the difference equation (3.9) at time k subject to the initial condition z at time l and input control sequence $\{u_j\}_{j=l}^{k-1}$. We assume the control and state sequences satisfy the constraints

$$u_k \in \mathbb{U}_k, \qquad x_k \in \mathbb{X}_k,$$

for all $k \geq 0$, where the sets $\mathbb{U}_k \subseteq \mathbb{R}^m$ and $\mathbb{X}_k \subseteq \mathbb{R}^n$ are closed and contain the origin.

Proceeding in an informal manner, we formulate RHC, given the measured state x_T, as the solution to the following optimal control problem

$$P_4(x_T): \qquad V_T^*(x_T) = \min_{\{u_k\}_{k=T}^{T+N-1}} \{V_T(\{u_k\}, x_T) \ : \ \{u_k\} \in \mathcal{U}_T^N(x_T)\}$$

where

$$V_T(\{u_k\}, x_T) := \sum_{k=T}^{T+N-1} l_k(x_k, u_k) + F_{T+N}(x_{T+N}),$$

with $x_k := x(k; x_T, T, \{u_j\})$ and $\mathcal{U}_T^N(z)$ is the set of feasible control sequences satisfying the input and state constraints subject to the initial condition $x_T = z$. We assume the stage costs $l_k : \mathbb{X}_k \times \mathbb{U}_k \to \mathbb{R}_{\geq 0}$ and the terminal penalty $F_k : \mathbb{X}_k \to \mathbb{R}_{\geq 0}$ for all $k \geq 0$. Let $\{u_{k|T}^*(x_T)\}_{k=0}^{N-1}$ denote the solution to $P_4(x_T)$. The feedback law is given by $u_T(\cdot) = u_{0|T}^*(\cdot)$.

If we assume the system (3.9) is (constrained) stabilizable and the stage penalties $l_k(\cdot)$ and $F_k(\cdot)$ satisfy certain technical conditions such as coercivity and boundedness, then a sufficient condition for asymptotic stability is

$$V_{T+N}^* - V_T^* \leq 0. \tag{3.10}$$

One may satisfy the inequality (3.10) by choosing the terminal penalty $F_\tau(\cdot)$ equal to the infinite horizon cost to go:

$$F_\tau(z) = \min_{\{u_k\}_{k=\tau}^{\infty}} \left\{ \sum_{k=\tau}^{\infty} l_k(u_k, x_k) \ : \ \{u_k\} \in \mathcal{U}_\tau^\infty(z) \right\}.$$

With the exception of the unconstrained linear quadratic problem (i.e. LQR), we are unable to calculate the cost to go. One alternative is to choose the terminal cost $F_T(\cdot)$ such that it satisfies the inequality

$$\begin{aligned} F_T(z) \ &\geq \ V_T^*(z), & (3.11a)\\ &= \min_{\{u_k\}_{k=T}^{T+N-1}} \left\{ \sum_{k=T}^{T+N-1} l_k(u_k, x_k) + F_{T+N}(x_{T+N}) : \{u_k\} \in \mathcal{U}_T^N(z) \right\}. & (3.11b) \end{aligned}$$

Recall condition **C2** requires also that the approximate arrival cost $\hat{\mathcal{Z}}(\cdot)$ satisfy an inequality, a lower bound rather than an upper bound. Furthermore, just as $\hat{\mathcal{Z}}_T(\cdot) = \hat{\phi}_T$ is the trivial solution in MHE, the choice $F_T(x) = \infty$ for $x \neq 0$, which we interpret as the constraint $x_T = 0$, is the trivial solution in RHC.

One may interpret the duality as follows; RHC requires an upper bound to the backward dynamic programming solution, whereas MHE requires a lower bound to the forward dynamic programming solution. Furthermore, to guarantee stability, it is not necessary to satisfy the minimum. Rather, it suffices to satisfy the inequality

$$F_T(z) \geq \sum_{k=T}^{T+N-1} l_k(u_{k|\infty}, x_{k|\infty}) + F_{T+N}(x_{T+N|\infty}) \tag{3.12}$$

where the sequence $\{u_{k|\infty}, x_{k|\infty}\}_{k=T}^{\infty} \in l_2(\mathbb{R}^m \times \mathbb{R}^n) \cap \mathcal{U}_T^{\infty}(x_{T|\infty})$. Notice the similarity to the inequality (3.7c). Likewise, the challenge is to construct the sequence $\{u_{k|\infty}, x_{k|\infty}\}_{k=T}^{\infty}$.

One strategy to construct the sequence is to employ local analysis (c.f. (Parisini and Zoppoli 1995) and (Chen and Allgöwer 1998)). If we assume the function $f_k^c(\cdot)$ is sufficiently smooth and locally stabilizable about the origin and the sets X_k and U_k both contain a neighborhood of the origin uniformly for all $k \geq 0$ (i.e. $\exists \epsilon > 0$ such that $B_\epsilon^1 \subset X_k$ and $B_\epsilon^2 \subset U_k$ for all $k \geq 0$), then there exists a linear feedback law K_f and positive invariant set $\mathcal{X}_\infty$, typically the level set of the local Lyapunov function $\mathcal{L}(\cdot)$,

$$\mathcal{X}_\infty(T) = \left\{ z \ : \ \begin{array}{ll} x_c(k; z, T) \in \mathbb{X}_k, & k \geq T \\ \{x_c(k; z, T)\}_{k=T}^{\infty} \in l_2(\mathbb{R}^n) & \\ K_f x_c(k; z, T) \in \mathbb{U}, & k \geq T \\ \mathcal{L}(z) \leq \alpha & \end{array} \right\},$$

where $x_c(k; z, l)$ denotes the solution of the difference equation (3.9) at time k subject to the initial condition z at time l and the feedback law $u_k = K_f x_k$. For all $z \in \mathcal{X}_\infty(T)$, one obtains the sequence $\{u_{k|\infty}, x_{k|\infty}\}_{k=T}^{\infty}$ by construction. Typically, one chooses $F_T(\cdot)$ such that it satisfies the more stringent inequality

$$F_T(z) \geq l(z, K_f z) + F_{T+1}(f_T^c(z, K_f z)) \ \ \forall z \in \mathcal{X}_\infty(T).$$

This condition holds if $F_T(\cdot)$ is a control Lyapunov function in the neighborhood of the origin. If we assume the control problem is time-invariant, then the inequality (3.12) reduces to

$$F(z) \geq l(z, K_f z) + F(f^c(z, K_f z)) \ \ \forall z \in \mathcal{X}_\infty.$$

A host of strategies exist for satisfying this inequality in a neighborhood of the origin (c.f. (Mayne et al. 2000)). For example, as the state is converging to the origin in RHC, one can linearize about the origin and use standard local analysis to construct a terminal penalty satisfying the inequality (3.12) in a neighborhood of the origin. There is no fixed equilibrium state in estimation, so we are unable to employ local analysis. Consequently it is necessary to construct the sequence $\{x_{k|\infty}, w_{k|\infty}\}$ by solving $P_3(T)$.

One obtains a stabilizing control law by receding horizon implementation of the following optimal control problem

$$P_5(x_T): \qquad V_T^*(x_T) = \min_{\{u_k\}_{k=T}^{T+N-1}} \{V_T(\{u_k\}, x_T) \ : \ \{u_k\} \in \mathcal{U}_T^N(x_T)\}$$

subject to the terminal constraint

$$x(T+N; x_T, T, \{u_j\}) \in \mathcal{X}_\infty(T+N),$$

where, for all $T \geq N$,

$$F_T(z) \geq l(z, K_f z) + F_{T+1}(f_T^c(z, K_f z)) \ \ \forall z \in \mathcal{X}_\infty(T).$$

One might expect we can enforce the inequality (3.11) by scaling the terminal penalty $F_{T+N}(\cdot)$ using a strategy similar to the one proposed in Section 3.4. However, to enforce the bound in RHC, we need to scale the terminal penalty $F_{T+N}(\cdot)$ at time T only after seeing the optimal trajectory at time $T+N$. But, the optimal trajectory at time $T+N$ depends upon the scaling at time T, due to the transition from x_T to x_{T+N}. This interdependency precludes the possibility of applying the scaling strategy to RHC. We do not have this interdependency problem in MHE, because we look backward. Whereas in RHC, we look forward and anticipate the closed-loop response. However, by looking forward and anticipating the closed-loop response, we are able to construct a stabilizing terminal penalty locally in an invariant region centered about the origin.

We encounter difficulties also if we attempt to enforce (3.11) by solving directly for the terminal penalty $F_N(\cdot)$ online. In addition to the computational difficulties associated with solving a functional inequality online, a solution may not exist for some future time $T \geq N$. If we initially choose the initial terminal penalty $F_N(\cdot)$ less than the steady-state cost to go, assuming it exists, then the terminal penalty $F_{jN}(\cdot)$ decreases necessarily by the monotonicity of dynamic programs as the index j increases. At some future time T, we may be unable to satisfy the functional inequality as $F_T(\cdot)$ is less than the cost associated with the sum $\sum_{k=T}^{T+N-1} l_k(\cdot)$. Consequently, we solve dynamic programs backward in control rather than forward in order to avoid such problems. The terminal penalty $F_N(\cdot)$ depends on the choice of the terminal penalty $F_{2N}(\cdot)$, so it is preferable to calculate the terminal penalty $F_{2N}(\cdot)$ first. Likewise, with MHE, the approximate arrival cost $\hat{\mathcal{Z}}_{2N}(\cdot)$ depends on our choice of the approximate arrival cost $\hat{\mathcal{Z}}_N(\cdot)$ at time N, so it is preferable to work forward.

3.7 Conclusion

In this chapter we investigated MHE as an online optimization strategy for estimating the state of constrained discrete-time systems. We provided conditions on the approximate arrival cost sufficient to guarantee stability. We may interpret this condition as requiring that we do not add information not specified by the data. For constrained linear models with quadratic objectives, the Kalman filter covariance satisfies the conditions for the approximate arrival cost. This result holds regardless of whether constraints are present. For constrained nonlinear state estimation, where an algebraic representation of the approximate arrival cost satisfying the stability requirement is generally unavailable, we proposed two alternative strategies. Both strategies circumvent the stability requirement by generating a stable reference trajectory. Using the reference trajectory, both strategies are able to scale the approximate arrival costs online to guarantee stability.

One of the goals of this research is to develop implementable strategies for constrained nonlinear state estimation. To account for the practical difficulties associated with online optimization, we proposed two stable suboptimal strategies. Neither of these strategies require global optimization. Instead, a feasible cost decrease is sufficient for stability. Even though the properties of these strategies were developed for nominal application, they demonstrate global optimization is not necessary for stability. Rather, we may expect a stable estimator from local or suboptimal solutions.

The practical significance of MHE is the ability to incorporate constraints explicitly. This feature distinguishes MHE from other strategies such as extended Kalman filtering and output error linearization. Furthermore, if the estimation problem translates into a problem of form $P_1(T)$, then we believe MHE is a natural engineering approximation to the full information problem, because the structure of MHE is not dictated by stability, but rather by performance and practicality. Stability results if one judiciously approximates the past data.

3.8 Appendix

3.8.1 Proof of Lemma 3.3.2

Proof. Recall x_0 denote the initial condition of the system (3.4). By the Lipschitz continuity of $f_k(\cdot)$, we have the inequality

$$\|x(T;x_0,0) - \hat{x}_T\| \leq c_f^N \|x(T-N;x_0,0) - \hat{x}_{T-N|T-1}\| + \sum_{k=T-N}^{T-1} c_f^{T-k} \, \|\hat{w}_{k|T-1}\|. \tag{3.13}$$

Let $\hat{y}_{k|T-1} := h_k(\hat{x}_{k|T-1})$. If we utilize the inverse triangle inequality, we obtain the inequality

$$\begin{aligned}\sum_{k=T-N}^{T-1} \|\hat{v}_{k|T-1}\| &= \sum_{k=T-N}^{T-1} \|y(k;x_0,0) - \hat{y}_{k|T-1}\|, \\ &\geq \sum_{k=T-N}^{T-1} \|y(k;x_0,0) - \bar{y}_k\| - \sum_{k=T-N}^{T-1} \|\bar{y}_k - \hat{y}_{k|T-1}\|,\end{aligned}$$

where $\bar{y}_k := y(k-(T-N);\hat{x}_{T-N|T-1}, T-N)$. Rearranging the inequality and utilizing the observability condition, we obtain the inequality

$$\begin{aligned}\sum_{k=T-N}^{T-1} \|\hat{v}_{k|T-1}\| + \sum_{k=T-N}^{T-1} \|\bar{y}_k - \hat{y}_{k|T-1}\| &\geq \sum_{j=0}^{N-1} \|y(j;x_0,0),0) - \bar{y}_k)\| \\ &\geq \varphi(\|x(T-N;x_0,0) - \hat{x}_{T-N|T-1}\|),\end{aligned}$$

where $\varphi(\cdot)$ be specified by the uniform observability assumption. Using (3.13) and the Lipschitz continuity of h_k, we have

$$\sum_{k=T-N}^{T-1} \|\bar{y}_k - \hat{y}_{k|T-1}\| \leq \sum_{j=T-N}^{T-1} \sum_{k=T-N}^{j} c_h c_f^{j-k} \, \|\hat{w}_{k|T-1}\|.$$

Combining the inequalities and using Fact 3.2.2, we obtain

$$\begin{aligned}&\|x(T;x_0,0) - \hat{x}_T\| \leq \\ &c_f^N \, \varphi^{-1}\left(\sum_{k=T-N}^{T-1} \|\hat{v}_{k|T-1}\| + \sum_{\ell=T-N}^{T-1} \sum_{k=T-N}^{\ell} c_h c_f^{\ell-k} \, \|\hat{w}_{k|T-1}\|\right) + \\ &\sum_{k=T-N}^{T-1} c_f^{T-k} x \|\hat{w}_{k|T-1}\|, \\ &\leq c_f^N \, \varphi^{-1}\left(\sum_{k=T-N}^{T-1} d^* + \sum_{\ell=T-N}^{T-1} \sum_{k=T-N}^{\ell} c_h c_f^{\ell-k} d^*\right) + \\ &\sum_{k=T-N}^{T-1} \prod_{j=k+1}^{T-1} c_f^{\ell-k} d^*, \\ &\leq k\left(\eta^{-1}\left(\sum_{k=T-N}^{T-1} L_k(\hat{w}_k, \hat{v}_k)\right)\right),\end{aligned}$$

where $k(\cdot)$ is a K-function. The existence of $\eta^{-1}(\cdot)$ and $k(\cdot)$ follow from Facts 3.2.2 and 3.2.3. $\square$

3.8.2 Proof of Proposition 3.4.3

Proof. For $T \leq N$, existence is established by Proposition 3.3.1. Now consider $T > N$ and let

$$\hat{\phi}_T^1 = \sum_{k=T-N}^{T-1} L_k(w_{k|\infty}, v_{k|\infty}) + \hat{\mathcal{Z}}_{T-N}(x_{T-N|\infty})$$

denote the finite cost, by assumption **A2** and property **C1**, associated with the feasible sequence $x_{T-N|\infty}$ and $\{w_{k|\infty}\}_{k=T-N}^{T-1}$ specified in assumption **A3** . Consider the set

$$\Lambda = \left\{ z, \{w_k\}_{k=T-N}^{T-1} \ : \ (z, \{w_k\}) \in \Omega_T^N, \ \ \hat{\phi}_T(z, \{w_k\}) \leq \hat{\phi}_T^1 \right\}$$

A solution exists under the stated assumption (see the proof of Proposition 3.3.1), if the set Λ is bounded. Assumption **A2** guarantees the sequence $\{w_k, v_k\}_{k=T-N}^{T-1}$ is bounded: $\|w_k\| + \|v_k\| \leq 2\eta^{-1}(\hat{\phi}_T^1)$. We conclude by demonstrating z is bounded. If we employ the inverse triangle inequality, we obtain

$$\sum_{k=T-N}^{T-1} \|v_k\| = \sum_{T-N}^{T-1} \|y_k - \bar{y}_k\| \geq \sum_{k=T-N}^{T-1} \|\bar{y}_k - y_{k|\infty}\| - \|y_k - y_{k|\infty}\|,$$

where $\bar{y}_k := y(k; z, T-N, \{w_j\})$ and $y_{k|\infty} := y(k; x_{T-N|\infty}, T-N, \{w_{,|\infty}\})$. Rearranging the inequality, we obtain

$$\sum_{k=T-N}^{T-1} \|v_k\| + \|y_k - y_{k|\infty}\| = \sum_{k=T-N}^{T-1} \|v_k\| + \|v_{k|\infty}\| \geq \sum_{k=T-N}^{T-1} \|\bar{y}_k - y_{k|\infty}\|.$$

If we employ again the inverse triangle inequality, we obtain

$$\begin{gathered}
\sum_{k=T-N}^{T-1} \|\bar{y}_k - y_{k|\infty}\| \geq \\
\sum_{k=T-N}^{T-1} \|\bar{y}_k - y(k; x_{T-N|\infty}, T-N)\| - \|y(k; x_{T-n|\infty}, T-N) - y_{k|\infty}\| \geq \\
\sum_{k=T-N}^{T-1} \|y(k; z, T-N) - y(k; x_{T-N|\infty}, T-N)\| - \\
\left(\sum_{k=T=N}^{T-1} \|\bar{y}_k - y(k; z, T-N)\| + \|y_{k|\infty} - y(k; x_{T-N|\infty}, T-N)\| \right).
\end{gathered}$$

Rearranging the inequality and applying the observability assumption, we obtain the inequality

$$\begin{gathered}
\sum_{T-N}^{T-1} \|y_{k|\infty} - \bar{y}_k\| + \|\bar{y}_k - y(k; z, T-N)\| + \|y_{k|\infty} - y(k; x_{T-N|\infty}, T-N)\| \geq \\
\varphi(\|x_{T-N|\infty} - z\|).
\end{gathered}$$

The first quantity $\|y_{k|\infty} - \bar{y}_k\|$ is bounded, using the triangle inequality, by $\|v_k\| + \|v_{k|\infty}\|$ and, consequently, by $2N\eta^{-1}(\hat{\phi}^1)$. To show the last two quantities are bounded, we employ assumption **A0** to obtain the following inequality

$$\begin{aligned}
\|\bar{y}_k - y(k; z, T-N)\| &\leq c_h \sum_{j=T-N}^{k-1} c_f^{k-j} \|w_k\|, \\
&\leq c_h \sum_{j=T-N}^{k-1} c_f^{k-j} 2N\eta^{-1}(\hat{\phi}_T^1).
\end{aligned}$$

Likewise, we have the inequality

$$\|y_{k|\infty} - y(k; x_{T-N|\infty}, T-N)\| \le c_h \sum_{j=T-N}^{k-1} c_f^{k-j} \eta^{-1}(\hat{\phi}_T^1).$$

Consequently, the quantity $\|x_{T-N|\infty} - z\|$ is bounded, and existence follows as claimed. □

3.8.3 Proof of Lemma 3.4.6

Proof. The details of the proof can be found in the proof of Lemma 3.3.2. If we choose

$$\theta_2(\cdot) := \varphi^{-1}\left(\sum_{k=T-N}^{T-1} \eta^{-1}(\cdot) + \sum_{\ell=T-N}^{T-1} \sum_{k=T-N}^{\ell} c_h c_f^{\ell-k} \eta^{-1}(\cdot)\right)$$

and

$$\theta_1(\cdot) := c_f^N\ \theta_2(\cdot) + \sum_{k=T-N}^{T-1} c_f^{T-k} \eta^{-1}(\cdot),$$

then the lemma follows. The existence of the K-functions $\varphi^{-1}(\cdot)$ and $\eta^{-1}(\cdot)$ follow from Fact 3.2.2, and existence of the K-function $\theta_1(\cdot)$ and $\theta_2(\cdot)$ follows from Fact 3.2.3. □

3.8.4 Proof of Lemma 3.4.10

Proof. The existence of $\bar{w}_k$ is guaranteed by **A4**. By assumption **A2**, we have the inequality $L_k(\bar{w}_k, \hat{v}_k^0) \le \gamma(\|\bar{w}_k\| + \|\hat{v}_k^0\|)$. By the definition of the model (3.1), we have the inequality

$$\|x(k+1; x_0, 0) - \hat{x}_{k+1}^0\| + \|f_k(x(k; x_0, 0), 0) - f_k(\hat{x}_k^0, 0)\| \ge \|f(\hat{x}_k^0, \bar{w}_k) - f(\hat{x}_k^0, 0)\|.$$

By assumptions **A0** and **A4**, we have the inequality

$$\kappa\left(\|x(k+1; x_0, 0) - \hat{x}_{k+1}^0\| + c_f\|(x(k; x_0, 0)) - \hat{x}_k^0\|\right) \ge \|\bar{w}_k\|.$$

By assumption **A0**, we have the inequality

$$\|\hat{v}_k^0\| = \|h_k(x(k; x_0, 0)) - h_k(\hat{x}_k^0)\| \le c_h\|x(k; x_0, 0) - \hat{x}_k^0\|.$$

Let $e_k := x(k; x_0, 0) - \hat{x}_k^0$. Combining the above inequalities, we obtain

$$L_k(\bar{w}_k, \hat{v}_k^0) \le \gamma\left(\kappa(\|e_{k+1}\| + c_f\|e_k\|) + c_h\|e_k\|\right)$$

Using Lemma 3.4.6, we obtain the inequality

$$L_k(\bar{w}_k, \hat{v}_k^0) \le \gamma\left(\kappa(\theta_1(\psi_{k+1}^*) + c_f\theta_1(\psi_k^*)) + c_h\theta_1(\psi_k^*)\right).$$

If we set $\mu(\cdot) := \gamma\left(\kappa(\theta_1(\cdot) + c_f\theta_1(\cdot)) + c_h\theta_1(\cdot)\right)$, then, by Fact 3.2.3, the lemma follows as claimed. □

Chapter 4

Constrained Linear State Estimation[1]

4.1 Introduction

The Kalman filter is the standard choice for estimating the state of a linear system when the measurements are noisy and the process disturbances are unmeasured. One reason for the popularity of the Kalman filter is that it possesses many important theoretical properties such as stability. Often additional insight about the process is available in the form of inequality constraints. With the addition of inequality constraints, however, general recursive solutions such as Kalman filtering are unavailable. One strategy for determining an optimal state estimate is to reformulate the estimation problem as a quadratic program. This formulation allows for the natural addition of inequality constraints. While there exist many strategies to solve efficiently quadratic programs with the particular structure of the linear estimation problem (c.f. (Biegler 1998)), the problem grows without bound as we collect more measurements.

Building on the success of model predictive control, moving horizon estimation (MHE) has been suggested as a practical method to incorporate inequality constraints in estimation (c.f. (Muske, Rawlings and Lee 1993), (Muske and Rawlings 1995), (Robertson, Lee and Rawlings 1996), (Tyler 1997), and (Rao and Rawlings 1998*a*)). The basic strategy of MHE is to reformulate the estimation problem as a quadratic program using a moving, fixed-size estimation window. The fixed-size estimation window is necessary to bound the size of the quadratic program. Because only a subset of the data is considered, stability questions arise. In this chapter we discuss moving horizon approximations for constrained linear state estimation. The results, derived independently, are a special case of the results presented in Chapter 3.

4.2 Problem Statement

Let the system generating the data sequence $\{y_k\}$ be modeled by the following linear, time-invariant, discrete-time system

$$x_{k+1} = Ax_k + Gw_k, \tag{4.1a}$$
$$y_k = Cx_k + v_k, \tag{4.1b}$$

where it is known that the states and disturbances satisfy the following constraints

$$x_k \in \mathbb{X}, \qquad w_k \in \mathbb{W}, \qquad v_k \in \mathbb{V}.$$

[1] **This chapter was published in an abridged form as Rao, Rawlings and Lee (1999)**

We assume $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^p$, and $w_k \in \mathbb{R}^m$ and the sets $\mathbb{X}$, $\mathbb{W}$, and $\mathbb{V}$ are polyhedral and convex (i.e. $\mathbb{X} = \{x \,:\, Dx \leq d\}$) with $0 \in \mathbb{W}$ and $0 \in \mathbb{V}$. Let $x(k; z, \{w_j\})$ denote the solution of model (4.1) at time k subject to the initial condition z and disturbance sequence $\{w_j\}_{j=0}^{k-1}$:

$$x(k; z, \{w_j\}) := A^k z + \sum_{j=0}^{k-1} A^{k-j-1} G w_j.$$

We formulate the constrained linear state estimation problem as the solution to the following quadratic problem

$$\phi_T^* = \min_{x_0, \{w_k\}_{k=0}^{T-1}} \phi_T(x_0,\ \{w_k\}) \tag{4.2}$$

subject to constraints

$$x_k \in \mathbb{X}, \qquad w_k \in \mathbb{W}, \qquad v_k \in \mathbb{V}, \tag{4.3}$$

where the objective function is defined by

$$\phi_T(x_0,\ \{w_k\}) = \sum_{k=0}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + (x_0 - \hat{x}_0)' \Pi_0^{-1} (x_0 - \hat{x}_0),$$

$x_k := x(k; x_0, \{w_j\})$, and $v_k := y_k - Cx(k; x_0, \{w_j\})$. We assume the matrices Q, R, and Π_0 are symmetric positive definite. The pair $(\hat{x}_0, \Pi_0)$ summarizes the prior information at time $T = 0$ and is part of the data of the state estimation problem. We refer to this problem as the **full information estimator**, because we consider all of the available measurements. The solution to (4.2) at time T is the unique pair $(\hat{x}_{0|T-1}, \{\hat{w}_{k|T-1}\}_{k=0}^{T-1})$, and the optimal pair yields the state estimate $\{\hat{x}_{k|T-1}\}_{k=0}^{T-1}$, where

$$\hat{x}_{k|T-1} := x(k, \hat{x}_{0|T-1}, \{\hat{w}_k\}).$$

To simplify notation, let $\hat{x}_j := \hat{x}_{j|j-1}$, where $\hat{x}_{0|-1} := \hat{x}_0$.

4.3 Moving Horizon Approximation

Efficient strategies exist for solving the quadratic program (4.2). However, the problem size grows with time as the estimator processes more data. As a result, the problem complexity scales at least linearly with T. To make the estimation problem tractable, we need to bound the problem size. One strategy to reduce the problem (4.2) to a fixed dimension quadratic program is to employ a moving horizon approximation. The basic strategy of the moving horizon approximation is to consider explicitly a fixed amount of data, while approximately summarizing the old data not explicitly accounted for by the estimator. The key to preserving stability and performance is how one *approximately* summarizes the old data.

Consider again the full information problem (4.2). We can rearrange the objective function $\phi_T(\cdot)$ by breaking the time interval into two pieces: $t_1 = \{k : 0 \leq k \leq T-N-1\}$ and $t_2 = \{k : T-N \leq k \leq T-1\}$.

$$\phi_T\left(x_0, \{w_k\}_{k=0}^{T-1}\right) \;=\; \phi_{T-N}\left(x_0, \{w_k\}\right) + \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k.$$

By the Markov property of the system (4.1), the quantity

$$\sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k$$

depends implicitly through the model (4.1) only on the state x_{T-N} and the decision variables w_k in the second time interval t_2. Exploiting this relation using forward dynamic programming, we can establish the equivalence between a full information problem and an estimation problem with a fixed size estimation window.

Consider the reachable set of states at time T generated by a feasible initial condition x_0 and disturbance sequence $\{w_k\}_{k=0}^{T-1}$:

$$\mathcal{R}_T = \left\{ x(T; x_0, \{w_j\}) \ : \ \begin{array}{l} x_0 \in \mathbb{X}, \\ x(k; x_0, \{w_j\}) \in \mathbb{X} \text{ for } k = 0, \dots, T, \\ w_k \in \mathbb{W} \text{ for } k = 0, \dots, (T-1) \end{array} \right\}.$$

For $z \in \mathcal{R}_T$, we define the **arrival cost**[2] as

$$\theta_T(z) := \min_{x_0, \{w_k\}_{k=0}^{T-1}} \left\{ \phi_T\left(x_0, \{w_j\}\right) : x(T; x_0, \{w_k\}) = z \right\},$$

where the minimization is subject to the constraints (4.3). It follows that $\theta_0(z) = (z - \hat{x}_0)'\Pi_0^{-1}(z - \hat{x}_0)$. Arrival cost is a fundamental concept in MHE, because the following equivalence can be established simply using forward dynamic programming:

$$\min_{x_0, \{w_k\}_{k=0}^{T-1}} \phi_T\left(x_0, \ \{w_k\}\right) \equiv$$

$$\min_{z, \{w_k\}_{k=T-N}^{T-1}} \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + \theta_{T-N}(z),$$

where the minimizations are subject to the constraints (4.3), $x_k := x(k - T - N; z, \{w_j\})$, and $v_k := y_k - Cx(k - T - N; z, \{w_j\})$.

The arrival cost compactly summarizes the effect of the data $\{y_k\}_{k=0}^{T-N-1}$ on the state x_{T-N}, thereby allowing us to fix the dimension of the optimization. We can view arrival cost as the analogue to the **cost to go** in standard backward dynamic programming. Loosely speaking in probabilistic terms, the arrival cost generates the conditional density function $p(x_{T-N}|y_0, \dots, y_{T-N-1})$ and vice-versa: the arrival cost is proportional to the negative logarithm of the conditional density function $p(x_{T-N}|y_0, \dots, y_{T-N-1})$ [3]. Hence, we may view arrival cost as an equivalent statistic (Striebel 1965) for the conditional density function $p(x_{T-N}|y_0, \dots, y_{T-N-1})$.

If we are able to construct analytic expressions for the arrival cost, then it is possible to develop recursive estimators. One example is Kalman filtering. Consider the unconstrained estimation problem (4.2). If we use the Kalman filter covariance update formula (Jazwinski 1970)

$$\Pi_T = GQG' + A\Pi_{T-1}A' - A\Pi_{T-1}C'(R + C\Pi_{T-1}C')^{-1}C\Pi_{T-1}A', \tag{4.4}$$

subject to the initial condition Π_0, then, assuming the matrix Π_T is invertible, we can express the arrival cost explicitly as

$$\theta_T(z) = (z - \hat{x}_T)'\Pi_T^{-1}(z - \hat{x}_T) + \phi_T^*,$$

where $\hat{x}_T$ denotes the optimal estimate at time T given the measurements $\{y_k\}_{k=0}^{T-1}$ and ϕ_T^* denotes the

[2] Other researchers have used the term **cost to come** (c.f (Başar and Bernhard 1995)) or **cost to arrive** (c.f. (Verdu and Poor 1987)).

[3] For example, if the conditional density function is normally distributed (i.e. $p(x_{T-N}|y_0, \dots, y_{T-N-1}) \sim N(\hat{x}_{T-N}, \Pi_{T-N})$), then $-\log\left(p(x_{T-N}|y_0, \dots, y_{T-N-1})\right) \propto (x_{T-N} - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(x_{T-N} - \hat{x}_{T-N})$.

optimal cost at time T. From the preceding arguments, we have

$$\min_{x_0,\ \{w_k\}_{k=0}^{T-1}} \phi_T(x_0,\ \{w_k\}) \equiv$$

$$\min_{z,\ \{w_k\}_{k=T-N}^{T-1}} \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + (z - \hat{x}_{T-N})' \Pi_{T-N}^{-1} (z - \hat{x}_{T-N}) + \phi_{T-N}^{*}.$$

We can extract the Kalman filter by considering a horizon of $N = 1$. For this scenario, we have

$$\phi_T(z, w_{T-1}) = v_{T-1}' R^{-1} v_{T-1} + w_{T-1}' Q^{-1} w_{T-1} + (z - \hat{x}_{T-1})' \Pi_{T-1}^{-1} (z - \hat{x}_{T-1}).$$

Substituting in the model equation (4.1), evaluating the minimum with respect to w_{T-1} and x_{T-1}, and using some algebra, we obtain the well known result

$$\hat{x}_T = A\hat{x}_{T-1} + A\Pi_{T-1}C'(R + C\Pi_{T-1}C')^{-1}(y_T - CA\hat{x}_{T-1})$$

for the Kalman filter.

Unfortunately, for the constrained problem, we are unable to generate an analytic expression for the arrival cost. Inequality constraints make the problem combinatorial, so general analytic expressions for the arrival cost are unavailable. One reasonable solution then is to approximate the arrival cost for the constrained problem with the arrival cost for the unconstrained problem. This choice has the desirable property that when the inequality constraints are inactive, the approximation is exact. Because we consider an approximation of the arrival cost, stability questions arise: does a poor choice of an approximate arrival cost lead to instability? As demonstrated in Section 4.6.1, the answer is yes. Instability may result for some systems if the arrival cost is improperly approximated. In the next section, we discuss the details of the stability arguments. As we demonstrate, it is not necessary to generate explicitly an analytic expression for the arrival cost. Rather, as discussed in Rao and Rawlings (1998*a*), the approximate arrival cost needs only to satisfy an inequality.

We formulate MHE as the solution to the following quadratic program

$$\hat{\phi}_T^{*} = \min_{z,\ \{w_k\}_{k=T-N}^{T-1}} \hat{\phi}_T(z,\ \{w_k\}), \tag{4.5}$$

subject to the constraints (4.3) where the objective function is defined by

$$\hat{\phi}_T(z,\ \{w_k\}) :=$$

$$\sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + (z - \hat{x}_{T-N}^{\text{mh}})' \Pi_{T-N}^{-1} (z - \hat{x}_{T-N}^{\text{mh}}) + \hat{\phi}_{T-N}^{*},$$

$x_k := x(k-(T-N); z, \{w_j\})$ and $v_k := y_k - Cx(k-(T-N); z, \{w_j\})$. The MHE cost $\hat{\phi}_T^*$ approximates the full information cost ϕ_T^* by replacing the arrival cost $\theta_{T-N}(z)$ with the quadratic approximation $(z - \hat{x}_{T-N}^{\text{mh}})'\Pi_{T-N}^{-1}(z - \hat{x}_{T-N}^{\text{mh}}) + \hat{\phi}_{T-N}^*$. The pair $(\hat{x}_{T-N}^{\text{mh}}, \Pi_{T-N})$ summarizes the prior information at time $T-N$. The vector $\hat{x}_{T-N}^{\text{mh}}$ is the moving horizon state estimate at time $T-N$ and the matrix Π_{T-N} is the solution to (4.4) subject to the initial condition Π_0. For $T \leq N$, MHE is equivalent to the full information estimator: $\hat{\phi}_T(\cdot) = \phi_T(\cdot)$. We assume at this point that the matrix Π_{T-N} is invertible; conditions for nonsingularity are discussed later. The solution to (4.2) at time T is the unique pair $(z^*, \{\hat{w}_{k|T-1}^{\text{mh}}\}_{k=T-N}^{T-1})$, and the optimal pair yields the state estimate $\{\hat{x}_{k|T-1}^{\text{mh}}\}_{k=T-N}^{T-1}$, where

$$\hat{x}_{k|T-1}^{\text{mh}} := x(k-(T-N); z^*, \{\hat{w}_j^{\text{mh}}\}).$$

To simplify notation, let $\hat{x}_j^{\text{mh}} := \hat{x}_{j|j-1}^{\text{mh}}$, where $\hat{x}_{0|-1}^{\text{mh}} := \hat{x}_0$. This formulation of MHE was first proposed by Muske et al. (1993) and Robertson et al. (1996).

4.4 Stability Analysis

When the inequality constraints (4.3) are not present, the solution to the quadratic program (4.2) may be obtained analytically, yielding the Kalman filter. The relationship between least squares and the Kalman filter is well known (c.f. (Bryson and Frazier 1963) and (Rauch, Tung and Striebel 1965)). Even with the addition of constraints, the estimator enjoys analogous stability properties. In particular, the constrained estimator is stable in the sense of an observer. The following discussion of observer stability is premised on classical Lyapunov stability theory for dynamical systems. The concepts are completely analogous to their classical counterpart. To account for constraints, we have modified the definition of stability in an analogous manner to Keerthi and Gilbert (1988).

Definition 4.4.1 *The estimator is a* **asymptotically stable observer** *for the system*

$$x_{k+1} = Ax_k, \qquad y_k = Cx_k. \tag{4.6}$$

if for any $\epsilon > 0$ there corresponds a number $\delta > 0$ and a positive integer $\bar{T}$ such that if $\|x_0 - \hat{x}_0\| \leq \delta$ and $\hat{x}_0 \in \mathbb{X}$, then $\|\hat{x}_T - A^T x_0\| \leq \epsilon$ for all $T \geq \bar{T}$ and $\hat{x}_T \to A^T x_0$ as $T \to \infty$.

The implications of constraints on the estimator are more subtle than for the regulator. In particular, the estimator has no control over the evolution of the state of the system. A poor choice of constraints may prevent convergence to the true state of the system (4.6). For further discussion of constraints, see Chapters 2 and 3. One solution is to require that the evolution of the system (4.6) respects the constraint $\mathbb{X}$ (i.e. $A^k x_0 \in \mathbb{X}$ for $k \geq 0$). While this assumption is reasonable, the constraints need to satisfy only the following weaker assumption to prove stability .

I Suppose the system (4.6) with initial condition x_0 generates the data (i.e. $y_k = CA^k x_0$). We assume there exists $x_{0|\infty}$, $\{w_{k|\infty}\}_{k=0}^{\infty}$, and $\sigma > 0$ such that

$$\sum_{k=0}^{\infty} v'_{k|\infty} R^{-1} v_{k|\infty} + w'_{k|\infty} Q^{-1} w_{k|\infty} + (x_{0|\infty} - \hat{x}_0)' \Pi_0^{-1} (x_{0|\infty} - \hat{x}_0) \leq \sigma \|x_0 - \hat{x}_0\|^2$$

and

$$x_{k|\infty} \in \mathbb{X}, \qquad w_{k|\infty} \in \mathbb{W}, \qquad v_{k|\infty} \in \mathbb{V},$$

where $x_{k|\infty} := x(k; x_{0|\infty}, \{w_{j|\infty}\})$ and $v_{k|\infty} := y_k - Cx(k; x_{0|\infty}, \{w_{j|\infty}\})$.

It is straightforward to demonstrate assumption **I** is a weaker assumption: if we choose $\sigma = \lambda_{\max}(\Pi_0^{-1})$, then assumption **I** follows if we assume the evolution (4.6) respects the constraints $\mathbb{X}$. Recall, by assumption, $0 \in \mathbb{W}$ and $0 \in \mathbb{V}$.

Assumption **I** states that if we consider an infinite amount of data generated by the system (4.6), then there exists a feasible state and disturbance trajectory that yields bounded cost. Assumption **I** is also a sufficient condition for the existence of a solution to the quadratic programs (4.2) and (4.5). The upper bound σ is necessary to prove stability. Without this bound, we have no reference for constructing a Lyapunov function. Unlike regulation where we have a strictly monotone nonincreasing cost function that is bounded below by zero, we have a strictly monotone nondecreasing cost function in estimation that is not necessarily bounded above (e.g. consider the case when assumption **I** is violated). The role of σ is to provide this upper bound when constraints prevent the estimator from tracking the system perfectly. Otherwise, without constraints, we can readily generate the upper bound with $\sigma = \lambda_{\max}(\Pi_0^{-1})$ (i.e. the cost of tracking the system perfectly).

In order to demonstrate stability, we require the following lemma.

Lemma 4.4.2 *Suppose (C, A) is observable and $N \geq n$. If*

$$\sum_{k=T-N}^{T-1} \hat{v}'_{k|T-1} R^{-1} \hat{v}_{k|T-1} + \hat{w}'_{k|T-1} Q^{-1} \hat{w}_{k|T-1} \rightarrow 0,$$

then $\|\hat{x}_T - x_T\| \rightarrow 0$.

Proof. The proof is available in the Appendix 4.8.1. □

Before discussing the stability of the MHE, we first state the following stability result for the full information estimator. A preliminary version of the proposition was proved by Muske et al. (1993), where convergence was established. For consistency with the MHE results, we prove the convergence and stability in the following proposition under slightly different assumptions and arguments than those used by (Muske et al. 1993).

Proposition 4.4.3 *Suppose the matrices Q, R, and Π_0 are positive definite, (C, A) is observable, and assumption* **I** *holds. Then, the constrained full information estimator is an asymptotically stable observer for the system (4.6).*

Proof. We assume throughout the proof $T \geq n$ (let $\tilde{T} = n$). We first demonstrate convergence. We know an optimal solution exists to (4.2), because the problem is a convex quadratic program and the feasible region is not empty: $x_{0|\infty}$ and $\{w_{k|\infty}\}_{k=0}^{T-1}$ (Frank and Wolfe 1956). Hence, by optimality, we have that $\sigma\|x_0 - \hat{x}_0\|^2 \geq \phi_k^*$ for all k. Writing out the cost function explicitly, we regroup the optimal cost as follows

$$\begin{aligned}
\phi_T^* &= \sum_{k=0}^{T-1} \hat{v}'_{k|T-1} R^{-1} \hat{v}_{k|T-1} + \hat{w}'_{k|T-1} Q^{-1} \hat{w}_{k|T-1} + (\hat{x}_{0|T-1} - \hat{x}_0)' \Pi_0^{-1} (\hat{x}_{0|T-1} - \hat{x}_0), \\
&= \hat{v}'_{T-1|T-1} R^{-1} \hat{v}_{T-1|T-1} + \hat{w}'_{T-1|T-1} Q^{-1} \hat{w}_{T-1|T-1} + \\
&\quad \sum_{k=0}^{T-2} \hat{v}'_{k|T-1} R^{-1} \hat{v}_{k|T-1} + \hat{w}'_{k|T-1} Q^{-1} \hat{w}_{k|T-1} + (\hat{x}_{0|T-1} - \hat{x}_0)' \Pi_0^{-1} (\hat{x}_{0|T-1} - \hat{x}_0),
\end{aligned}$$

where $\hat{v}_{k|T-1} := y_k - C\hat{x}_{k|T-1}$. Since $\hat{x}_{0|T-1}$ and $\{\hat{w}_{k|T-1}\}_{k=0}^{T-2}$ are feasible at time index $T-1$, we obtain the inequality

$$\sum_{k=0}^{T-2} \hat{v}'_{k|T-1} R^{-1} \hat{v}_{k|T-1} + \hat{w}'_{k|T-1} Q^{-1} \hat{w}_{k|T-1} + (\hat{x}_{0|T-1} - \hat{x}_0)' \Pi_0^{-1} (\hat{x}_{0|T-1} - \hat{x}_0) \geq \phi_{T-1}^*$$

by optimality. This inequality implies

$$\phi_T^* - \phi_{T-1}^* \geq \hat{v}'_{T-1|T-1} R^{-1} \hat{v}_{T-1|T-1} + \hat{w}'_{T-1|T-1} Q^{-1} \hat{w}_{T-1|T-1} \geq 0.$$

Since ϕ_T^* is nondecreasing and bounded above by $\sigma\|x_0 - \hat{x}_0\|^2$, the sequence of optimal costs ϕ_T^* converges to $\phi_\infty^* \leq \sigma\|x_0 - \hat{x}_0\|^2 < \infty$. Convergence implies for some fixed $N \geq n$,

$$\phi_T^* - \phi_{T-N}^* \geq \sum_{k=T-N}^{T-1} \hat{v}'_{k|T-1} Q^{-1} \hat{v}_{k|T-1} + \hat{w}'_{k|T-1} Q^{-1} \hat{w}_{k|T-1} \rightarrow 0$$

as $T \rightarrow \infty$. By Lemma 4.4.2, it follows that the estimation error $\|\hat{x}_T - A^T x_0\| \rightarrow 0$ as $T \rightarrow \infty$.

To prove stability, let $\epsilon > 0$ and choose $\varrho > 0$ sufficiently small for $T = n$ as specified by Lemma 4.4.2. If we choose $\delta > 0$ such that $\sigma\delta^2 < \varrho$, then we obtain the following inequality for all

$T \geq n$.

$$\begin{aligned} \sigma\delta^2 &\geq \sum_{k=0}^{T-1} \hat{v}'_{k|T-1} R^{-1} \hat{v}_{k|T-1} + \hat{w}'_{k|T-1} Q^{-1} \hat{w}_{k|T-1} + (\hat{x}_{0|T-1} - \hat{x}_0)' \Pi_0^{-1} (\hat{x}_{0|T-1} - \hat{x}_0), \\ &\geq \sum_{k=T-n}^{T-1} \hat{v}'_{k|T-1} R^{-1} \hat{v}_{k|T-1} + \hat{w}'_{k|T-1} Q^{-1} \hat{w}_{k|T-1}. \end{aligned}$$

Hence, if the initial estimation error $\|x_0 - \hat{x}_0\| \leq \delta$, then the estimation error $\|\hat{x}_T - A^T x_0\| \leq \epsilon$ for all $T \geq \tilde{T} = n$ as claimed. □

To establish asymptotic stability for MHE, we require the following lemma.

Lemma 4.4.4 *The Kalman filter covariance matrix Π_T satisfies the following inequality for all $p \in \mathcal{R}_T$*

$$\begin{aligned} (p - \hat{x}_T^{mh})' \Pi_T^{-1} (p - \hat{x}_T^{mh}) + \hat{\phi}_T^* &\leq \min_{z,\, \{w_k\}_{k=T-N}^{T-1}} \left\{ \hat{\phi}_T(z,\ \{w_k\}) \ :\ x(N; z, \{w_j\}) = p \right\}, \\ &:= \hat{\theta}_T(p), \end{aligned}$$

where the minimization is subject to the constraints (4.3).

Proof. The proof is available in Appendix 4.8.2. □

Before we establish stability, we need to characterize conditions that guarantee the matrix Π_T is positive definite (invertible). If we assume that (C, A) is detectable and $(A, GQ^{-1/2})$ is controllable, then

$$\lim_{T \to \infty} \Pi_T = \Pi_\infty,$$

where $\Pi_\infty > 0$ is the unique steady-state solution to the Riccati equation (4.4) (de Souza, Gevers and Goodwin 1986). If we choose $\Pi_0 \geq \Pi_\infty$, then Π_k is positive definite for all $k \geq 0$ (Bitmead, Gevers, Petersen and Kaye 1985). As an alternative, if the matrix G is nonsingular (in which case GQG^T is positive definite), then Π_k is also positive definite for all $k \geq 0$.

Proposition 4.4.5 *Suppose the matrices Q, R, and Π_0 are positive definite, (C, A) is observable, assumption* **I** *holds, $N \geq n$, and either*

1. *The matrix G is nonsingular, or*
2. *$(A, GQ^{-1/2})$ is controllable and $\Pi_0 \geq \Pi_\infty$.*

Then the constrained moving horizon estimator is an asymptotically stable observer for the system (4.6).

Proof. We begin by demonstrating convergence. An optimal solution to (4.5) exists (Frank and Wolfe 1956), because the problem (4.5) is a convex quadratic program and the feasible region is not empty: the pair $x_{T-N|\infty}$ and $\{w_{k|\infty}\}_{k=T-N}^{T-1}$ is feasible. By definition,

$$\hat{\phi}_T^* - \hat{\phi}_{T-N}^* \geq \sum_{k=T-N}^{T-1} \hat{v}_{k|T-1}^{\mathrm{mh}'} R^{-1} \hat{v}_{k|T-1}^{\mathrm{mh}} + \hat{w}_{k|T-1}^{\mathrm{mh}'} Q^{-1} \hat{w}_{k|T-1}^{\mathrm{mh}},$$

where $\hat{v}_{k|T-1}^{\mathrm{mh}} := y_k - C\hat{x}_{k|T-1}^{\mathrm{mh}}$. To demonstrate $\sigma\|x_0 - \hat{x}_0\|^2$ is a uniform bound, we proceed using an induction argument. For $T \leq N$, we have by optimality

$$\hat{\phi}_T^* \leq \theta_T(x_{T|\infty}) \leq \sigma\|x_0 - \hat{x}_0^{\mathrm{mh}}\|^2.$$

For $T \geq N$, Lemma 4.4.4 guarantees

$$\theta(x_{T|\infty}) = \hat{\theta}(x_{T|\infty}) \geq (x_{T|\infty} - \hat{x}_T)'\Pi_T^{-1}(x_{T|\infty} - \hat{x}_T) + \hat{\phi}_T^*.$$

Let us now assume, for $T > N$,

$$\theta(x_{T-N|\infty}) \geq (x_{T-N|\infty} - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(x_{T-N|\infty} - \hat{x}_{T-N}) + \hat{\phi}_{T-N}^*.$$

for the induction argument. Utilizing the optimality principle, the induction assumption, and properties of the arrival cost, for all $T \geq N$,

$$\begin{aligned}
\sigma\|x_0 - \hat{x}_0\|^2 \quad & \geq \quad \min_{z,\,\{w_k\}_{k=T-N}^{T-1}} \left\{ \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + \hat{\theta}_{T-N}(z) \ : x(N; z, \{w_j\}) = x_{T|\infty} \right\}, \\
& \qquad \text{(by optimality and assumption } \mathbf{I} \text{)} \\
& \geq \quad \min_{z,\,\{w_k\}_{k=T-N}^{T-1}} \left\{ \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + \right. \\
& \qquad \left. (z - \hat{x}_{T-N}^{\mathrm{mh}})'\Pi_{T-N}^{-1}(z - \hat{x}_{T-N}^{\mathrm{mh}}) + \hat{\phi}_{T-N}^* \ : \ x(N; z, \{w_j\}) = x_{T|\infty} \right\}, \\
& \qquad \text{(by the induction assumption)} \\
& \geq \quad (x_{T|\infty} - \hat{x}_T^{\mathrm{mh}})'\Pi_T^{-1}(x_{T|\infty} - \hat{x}_T^{\mathrm{mh}}) + \hat{\phi}_T^* \\
& \qquad \text{(by Lemma 4.4.4)} \\
& \geq \quad \hat{\phi}_T^*,
\end{aligned}$$

where both minimizations are subject to the constraints (4.3). Hence, the sequence $\{\hat{\phi}_T^*\}$ is a monotone nondecreasing and bounded above by $\sigma\|x_0 - \hat{x}_0^{\mathrm{mh}}\|^2$. Convergence implies

$$\sum_{k=T-N}^{T-1} \hat{v}_{k|T-1}^{\mathrm{mh}'} R^{-1} \hat{v}_{k|T-1}^{\mathrm{mh}} + \hat{w}_{k|T-1}^{\mathrm{mh}'} Q^{-1} \hat{w}_{k|T-1}^{\mathrm{mh}} \to 0$$

as $T \to \infty$. Lemma 4.4.2 guarantees the estimation error $\|\hat{x}_T^{\mathrm{mh}} - A^T x_0\| \to 0$ as $T \to \infty$.

To prove stability, let $\epsilon > 0$ and choose $\varrho > 0$ sufficiently small for $T = N$ as specified by Lemma 4.4.2. If we choose $\delta > 0$ such that $\sigma\delta^2 < \varrho$, then we obtain the following inequality for all $T \geq N$.

$$\begin{aligned}
\sigma\delta^2 \quad & \geq \quad \sum_{k=T-N}^{T-1} \hat{v}_{k|T-1}^{\mathrm{mh}'} R^{-1} \hat{v}_{k|T-1}^{\mathrm{mh}} + \hat{w}_{k|T-1}^{\mathrm{mh}'} Q^{-1} \hat{w}_{k|T-1}^{\mathrm{mh}} \\
& \qquad + (\hat{x}_{T-N|T-1}^{\mathrm{mh}} - \hat{x}_{T-N}^{\mathrm{mh}})'\Pi_{T-N}^{-1}(\hat{x}_{T-N|T-1}^{\mathrm{mh}} - \hat{x}_{T-N}^{\mathrm{mh}}) + \hat{\phi}_{T-N}^*, \\
& \geq \quad \sum_{k=T-N}^{T-1} \hat{v}_{k|T-1}^{\mathrm{mh}'} R^{-1} \hat{v}_{k|T-1}^{\mathrm{mh}} + \hat{w}_{k|T-1}^{\mathrm{mh}'} Q^{-1} \hat{w}_{k|T-1}^{\mathrm{mh}},
\end{aligned}$$

Hence, if the initial estimation error $\|x_0 - \hat{x}_0^{\mathrm{mh}}\| \leq \delta$, then the estimation error $\|\hat{x}_T^{\mathrm{mh}} - A^T x_0\| \leq \epsilon$ for all $T \geq \tilde{T} = N$ as claimed. □

Remark 4.4.6 *When inequality constraints are not included, MHE is equivalent to the Kalman filter. Proposition 4.4.5, therefore, establishes that the Kalman filter is stable under the stated conditions.*

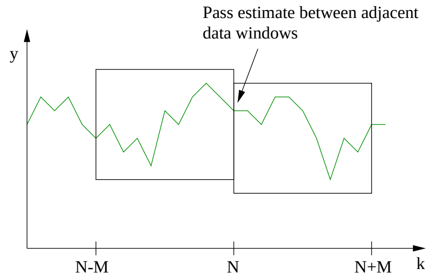


Figure 4.1: A diagram of the filter update strategy for passing information forward in time.

We may also formulate the constrained steady-state MHE where the objective function is now defined as

$$\hat{\phi}_T^{\infty}(z,\ \{w_k\}) := \sum_{k=T-1}^{T-N} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + (z - \hat{x}_{T-N})' \Pi_{\infty}^{-1} (z - \hat{x}_{T-N}) + \hat{\phi}_{T-N}^{*}.$$

For $T \leq N$, we choose $\hat{\phi}_T^{\infty}(\cdot) = \phi_T(\cdot)$ with $\Pi_0 = \Pi_{\infty}$. Demonstrating the stability of steady-state MHE is immediate. In Proposition 4.4.5, we proved stability for all $\Pi_0 > 0$. If we choose $\Pi_0 = \Pi_{\infty}$, then $\Pi_T = \Pi_{\infty}$ for all T. We state this result as the following corollary to Proposition 4.4.5.

Corollary 4.4.7 *Suppose the matrices Q and R are positive definite, (C, A) is observable, $(A, GQ^{-1/2})$ is controllable, assumption* **I** *holds, and $N \geq n$. Then the constrained state-state moving horizon estimator is an asymptotically stable observer for the system*

4.5 Smoothing Update

In our development of the MHE, we use a filter update to summarize the past information. With the filter update we transfer the prior information to current estimate window by conditioning the estimates at time T using $\hat{x}_{T-N}$. The conditioning is the result of the approximate arrival cost

$$(x_{T-N} - \hat{x}_{T-N})' \Pi_{T-N}^{-1} (x_{T-N} - \hat{x}_{T-N})$$

achieving its minimum at $\hat{x}_{T-N}$. A schematic of the filter update strategy is shown in Figure 4.1. We can interpret the filter update strategy as follows. If we consider the conditional probability density of the state sequence $\{x_k\}_{k=0}^{T}$ given the data sequence $\{y_k\}_{k=0}^{T-1}$, we have

$$
\begin{aligned}
p(x_{T-N},\dots,x_T|y_0,\dots,y_{T-1}) &= \frac{p(x_{T-N},\dots,x_T,y_0,\dots,y_{T-1})}{p(y_0,\dots,y_{T-1})} \\
&\quad (p(b|a) = p(a,b)/p(a)) \\
&= p(x_{T-N+1},\dots x_T, y_{T-N},\dots,y_{T-1}|x_{T-N},y_0,\dots,y_{T-N-1}) \times \\
&\quad p(x_{T-N}|y_0,\dots y_{T-N-1})\frac{p(y_0,\dots y_{T-N-1})}{p(y_0,\dots,y_{T-1})}, \\
&\quad (p(a,b,c,d) = p(ab|cd)p(c|d)p(d)) \\
&= p(x_{T-N+1},\dots x_T, y_{T-N},\dots,y_{T-1}|x_{T-N}) \times \\
&\quad p(x_{T-N}|y_0,\dots y_{T-N-1})\frac{p(y_0,\dots y_{T-N-1})}{p(y_0,\dots,y_{T-1})}, \\
&\quad (\text{the Markov property}) \\
&= p(y_{T-N},\dots,y_{T-1}|x_{T-N},\dots,x_{T-1}) \times \\
&\quad p(x_{T-N+1},\dots x_T|x_{T-N}) \times \\
&\quad p(x_{T-N}|y_0,\dots y_{T-N-1})\frac{p(y_0,\dots y_{T-N-1})}{p(y_0,\dots,y_{T-1})}. \\
&\quad (p(a,b|c) = p(a|b,c)p(b|c) \text{ and the Markov property})
\end{aligned}
$$

Because the probability densities $p(y_0,\dots y_{T-N-1})$ and $p(y_0,\dots,y_{T-1})$ are independent of the choice of the state sequence $\{x_k\}_{k=T-N}^{T}$, we obtain

$$
\begin{aligned}
p(x_{T-N},\dots,x_T|y_0,\dots,y_{T-1}) &\propto p(y_{T-N},\dots,y_{T-1}|x_{T-N},\dots,x_{T-1}) \times \\
&\quad p(x_{T-N+1},\dots x_T|x_{T-N}) \times \\
&\quad p(x_{T-N}|y_0,\dots y_{T-N-1}), \\
&\quad (\text{the Markov property and the model structure}) \\
&\propto \prod_{k=T-N}^{T-1} p(y_k|x_k)p(x_{k+1}|x_k) \times \\
&\quad p(x_{T-N}|y_0,\dots y_{T-N}).
\end{aligned}
$$

We may view the term $\prod_{k=T-N}^{T-1} p(y_k|x_k)p(x_{k+1}|x_k)$ as the moving horizon contribution to the state estimate and the term $p(x_{T-N}|y_0,\dots y_{T-N-1})$ as the contribution of the initial penalty, or arrival cost, to the estimate. The initial penalty approximates an equivalent statistic by summarizing the past information not included in the current estimation window.

Rather than condition the estimate at time T on $\hat{x}_{T-N}$, we may also condition the estimate on $\hat{x}_{T-N|T-2}$. With the filter update, we ignore the influence of the data $\{y_k\}_{T-N}^{T-2}$ on our knowledge of x_{T-N}. A diagram of the smoothing update strategy is shown in Figure 4.2. This problem was first studied by Findeisen (1997).

In analogous manner to the filter update, we have a probabilistic interpretation for MHE with a smoothing update. We begin our derivation by first considering the information transition from time

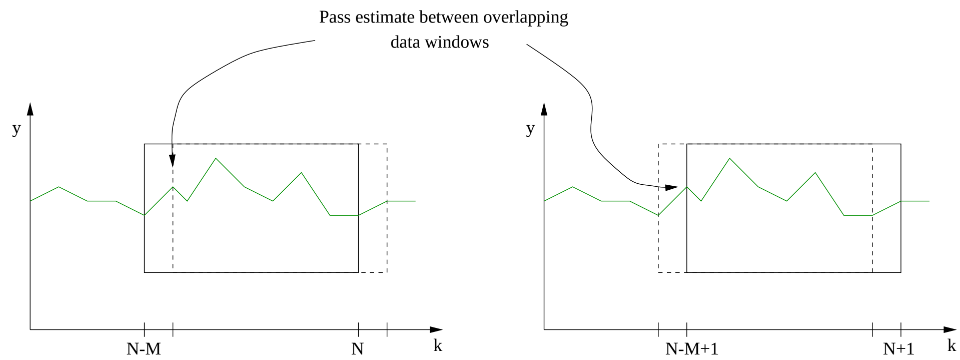


Figure 4.2: A diagram of the smoothing strategy for passing information forward in time.

$T-1$ to T. Utilizing the properties of conditional densities, we obtain

$$\begin{aligned}
p(x_{T-N},\dots,x_T|y_0,\dots,y_{T-1}) &= p(x_T,y_{T-1}|x_{T-N},\dots,x_{T-1},y_0,\dots,y_{T-2}) \times \\
&\quad p(x_{T-N},\dots,x_{T-1}|y_0,\dots,y_{T-2}) \times \\
&\quad \frac{p(y_0,\dots,y_{T-2})}{p(y_0,\dots,y_{T-1})}, \\
&\quad (p(a,b|c,d) = p(bd|ac)(a|c)p(c)/p(d)) \\
&= p(x_T,y_{T-1}|x_{T-N},\dots,x_{T-1}) \times \\
&\quad p(x_{T-N},\dots,x_{T-1}|y_0,\dots,y_{T-2}) \times \\
&\quad \frac{p(y_0,\dots,y_{T-2})}{p(y_0,\dots,y_{T-1})}, \\
&\quad \text{(the Markov property)} \\
&= p(y_{T-1}|x_{T-1})p(x_T|x_{T-1}) \times \\
&\quad p(x_{T-N},\dots,x_{T-1}|y_0,\dots,y_{T-2}) \times \\
&\quad \frac{p(y_0,\dots,y_{T-2})}{p(y_0,\dots,y_{T-1})}. \\
&\quad \text{(the Markov property)}
\end{aligned}$$

From our solution at time $T-1$, we possess information concerning

$$p(x_{T-N},\dots,x_{T-1}|y_0,\dots,y_{T-2}).$$

However, we prefer to retain the explicit use of the data window in the estimation procedure. In particular, we reformulate the conditional density as follows

$$\begin{aligned}
p(x_{T-N},\dots,x_{T-1}|y_0,\dots,y_{T-2}) &= p(x_{T-N+1},\dots,x_{T-1}|x_{T-N},y_0,\dots,y_{T-2}) \times \\
&\quad p(x_{T-N}|y_0,\dots,y_{T-2}).
\end{aligned}$$

We obtain the following expression

$$
\begin{aligned}
&p(x_{T-N+1},\dots,x_{T-1}|x_{T-N},y_{T-N},\dots,y_{T-2})\\
&\quad= \frac{p(x_{T-N+1},\dots,x_{T-1},y_{T-N},\dots y_{T-2}|x_{T-N})}{p(y_{T-N},\dots,y_T|x_{T-N})},\\
&\qquad (p(a,b|c)=p(a|bc)p(b|c))\\
&\quad= p(y_{T-N},\dots y_{T-2}|x_{T-N},\dots,x_{T-1})\frac{p(x_{T-N+1},\dots,x_{T-1}|x_{T-N})}{p(y_{T-N},\dots,y_T|x_{T-N})}.\\
&\qquad \text{(the Markov property)}
\end{aligned}
$$

Hence, we obtain

$$
\begin{aligned}
p(x_{T-N},\dots,x_T|y_0,\dots,y_{T-1}) \quad \propto \quad & \prod_{k=T-N}^{T-1} p(y_k|x_k)p(x_{k+1}|x_k)\times\\
& \frac{p(x_{T-N}|y_0,\dots,y_{T-2})}{p(y_{T-N},\dots,y_{T-2}|x_{T-N})}.
\end{aligned}
$$

We may view the first term as the moving horizon contribution to the estimate and the second term as the contribution of the initial penalty to the estimate. Here, the initial penalty is the smoothed equivalent statistic.

Before discussing the arrival cost using the smoothing update strategy, we first derive an expression for

$$
\frac{p(x_{T-N}|y_0,\dots,y_{T-2})}{p(y_{T-N},\dots,y_{T-2}|x_{T-N})}.
$$

For linear unconstrained systems we can calculate the smoothed equivalent statistic as follows. The numerator is the smoothed covariance of the state estimate. Rauch et al. (1965) demonstrate

$$
p(x_k|y_0\dots y_T)\left\{:=p(x_{k|T})\right\}\sim N(\hat{x}_{k|T},\Pi_{k|T}),
$$

where the smoothed covariance is obtained from the following backward Riccati equation

$$
\Pi_{k|T}=\Pi_{k|k}+\Pi_{k|k}A_k'\Pi_{k+1|T}^{-1}(\Pi_{k+1|T}-\Pi_{k+1|k})\Pi_{k+1|T}^{-1}A_k\Pi_{k|k}, \tag{4.7}
$$

where

$$
\Pi_{k|k}=\Pi_k-\Pi_kC'\left(R+C\Pi_kC'\right)^{-1}C\Pi_k
$$

and $\Pi_{k|k-1}:=\Pi_k$. To evaluate the denominator, which prevents the estimator from using the data $\{y_k\}_{k=T-N}^{T-2}$ twice (i.e. accounting for overlap between the estimation windows), we have the following expression

$$
p(y_{T-N},\dots,y_{T-2}|x_{T-N})\sim N\left(\mathcal{O}_{N-2}x_{T-N},W_{N-2}\right),
$$

where the expressions for $\mathcal{O}_{N-2}$ and W_{N-2} are given in Appendix 4.8.3. The derivation follows from the elementary properties of linear Gaussian difference equations.

We formulate MHE with the smoothing update by using the objective function

$$
\hat{\phi}_T(z,\{w_k\})=\sum_{k=T-N}^{T-1} w_k'Q^{-1}w_k+v_k'R^{-1}v_k+\Gamma_{T-N}(z)+\hat{\phi}_{T-N}^*,
$$

where

$$\Gamma_{T-N}(z) = (z - \hat{x}_{T-N|T-2})'\Pi^{-1}_{T-N|T-2}(z - \hat{x}_{T-N|T-2}) - \left(\mathcal{Y}_T^{N-2} - \mathcal{O}_{N-2}z\right)' W_{N-2}^{-1} \left(\mathcal{Y}_T^{N-2} - \mathcal{O}_{N-2}z\right) + (\hat{\phi}^*_{T-1} - \hat{\phi}^*_{T-N}),$$

and

$$\mathcal{Y}_T^{N-2} := \begin{bmatrix} y'_{T-N}, & y'_{T-N-1}, & \dots, & y'_{T-2} \end{bmatrix}'.$$

Expressions for $\mathcal{O}_{N-2}$ and W_{N-2} are given in Appendix 4.8.3.

To prove stability, it suffices in light of Proposition 4.4.5 to demonstrate that $\Gamma_T(\cdot)$ satisfies the inequality in Lemma 4.4.4.

Lemma 4.5.1 *Suppose the matrix $\Pi_{j|T-1}$ is positive definite. Then, we have for $j < T$*

$$(z - \hat{x}_{j|T-1})'\Pi^{-1}_{j|T-1}(z - \hat{x}_{j|T-1}) + \phi^*_T = \min_{x_0, \{w_k\}_{k=0}^{T-1}} \{\phi_T(x_0, \{w_k\}) \ : \ x(j; x_0, \{w_k\} = z\}. \tag{4.8}$$

Proof. This equality follows from the smoothing results for linear discrete-time system (c.f. (Rauch et al. 1965) and (Bryson and Ho 1975)). □

Lemma 4.5.2 *Suppose the matrix $\Pi_{T-N|T-2}$ is positive definite. Then, for all $p \in \mathcal{R}_T$ and $j < T$*

$$\begin{aligned} \Gamma_T(p) + \hat{\phi}^*_T \quad &\leq \quad \min_{z, \{w_k\}_{k=T-N}^{T-1}} \left\{ \hat{\phi}_T(z, \{w_k\}) \ : \ x(N; z, \{w_k\}) = p \right\}, \\ &:= \quad \hat{\theta}_T(p). \end{aligned}$$

where the minimization is subject to the constraints (4.3).

Proof. The proof is available in the Appendix 4.8.4. □

Corollary 4.5.3 *Suppose the matrices Q, R, Π_0, and $\Pi_{T-N|T-2}$ for all $T \geq N$ are positive definite, (C, A) is observable, assumption* **I** *holds, $N \geq n$. Then the constrained moving horizon estimator with a smoothing update is an asymptotically stable observer for the system (4.6).*

4.6 Example of inequality constraints yielding improved estimates.

Consider the following discrete-time system[4]

$$x_{k+1} = \begin{bmatrix} 0.9962 & 0.1949 \\ -0.1949 & 0.3815 \end{bmatrix} x_k + \begin{bmatrix} 0.03393 \\ 0.1949 \end{bmatrix} w_k, \qquad y_k = \begin{bmatrix} 1 & -3 \end{bmatrix} + v_k. \tag{4.9}$$

We assume $\{v_k\}$ is sequence of independent, zero mean, normally distributed random variables with covariance 0.01, and $w_k = |z_k|$ where $\{z_k\}$ is a sequence of independent, zero mean, normally distributed random variables with unit covariance. We also assume the initial state x_0 is normally distributed with zero mean and covariance equal to the identity.

We formulate the constrained estimation problem with $Q = 1$, $R = 0.01$, $\Pi_0 = 1$, and $\hat{x}_0 = 0$. For the MHE, we choose $N = 10$. To capture our knowledge of the random sequence w_k, we add the

[4]This state space system is a realization of the following system $G(s) = \frac{-3s+1}{s^2+3s+1}$ sampled with a zero-order hold and sampling time of 0.3.

inequality constraint $w_k \geq 0$. Note, this formulation yields the *optimal* Bayesian estimate. A comparison of the Kalman filter, full information estimator, and MHE with a filter update for a single realization of (4.9) is shown in Figure 4.3. As expected, the performance of the constrained estimators is superior to the Kalman filter, because the constrained estimators possess, with the addition of the inequality constraints, the proper statistics of the disturbance sequence w_k. Hence, the constrained estimation problem formulated above accurately models the random variable w_k.

If we consider the statistics of the random variable w_k, it is important to note that the mean is not zero and the covariance is not 1. Rather, the mean is $2/\sqrt{2\pi}$ and the covariance is $(1 - 2/\pi)$. When we consider the negative inverse logarithm of the probability density function, however, we have

$$-\log p_{w_k}(w_k) \propto \frac{1}{2} w_k' w_k \qquad \text{for } w_k \geq 0.$$

Note, therefore, that constraints allow for non-Gaussian disturbances.

4.6.1 Example of instability due to a poor choice of the initial penalty.

Consider the problem of estimating a unit step disturbance to the system (4.9) with initial condition $x_0 = 0$ when there is no process or measurement noise (i.e. a deterministic estimation problem). We model the disturbance as an integrator with the following process model:

$$x_{k+1} = \begin{bmatrix} 0.9962 & 0.1949 & 0.03393 \\ -0.1949 & 0.3815 & 0.1949 \\ 0 & 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} w_k, \qquad y_k = \begin{bmatrix} 1 & -3 & 0 \end{bmatrix} + v_k. \tag{4.10}$$

The state $x^{(3)}$ represents our estimate of the disturbance. We formulate the unconstrained MHE with $Q = 1$, $R = 1$, $N = 10$, and $\hat{x}_0 = 0$. For this example, we consider only steady-state MHE. We choose the initial penalty

$$\frac{1}{\alpha}(z - \hat{x}_{j|j-1})' \Pi_\infty^{-1} (z - \hat{x}_{j|j-1}).$$

Figure 4.4 shows the response of the MHE for different values of α. When $\alpha = 1$, MHE is equivalent to the steady-state Kalman filter, because no constraints are present. If we choose α sufficiently small, the estimator diverges, because the past data is weighted too strongly and the estimator is unable to "keep up" with the data. The stability limit for α in this example is approximately 0.23. For many system (i.e. asymptotically stable), the estimator is stable even when $\alpha \to 0$. For this example, the non-minimum phase behavior and the addition of an integrator were necessary to induce instability.

4.6.2 Examples of instability due to infeasible inequality constraints.

It is simple to demonstrate that an empty feasible region causes estimator divergence. For example, consider the system with $A = 1$, $C = 1$, and the initial condition $x_0 = 0$. If we choose the constraints such that $|x_k| \geq 1$, then it is impossible to converge to the true state of the system without violating the constraints. For unstable systems, one might expect poorly chosen constraints on w_k result also in instability: if we choose the constraints on w_k too rigidly, then the unstable modes of the system dominate Gw_k. If there are no constraints on state and $0 \in \mathbb{W}$, then this scenario does not occur. The reason is that we are always free to choose x_0 equal to the true state of the system. This choice provides an upper bound to the optimal cost for the estimator. Therefore, the estimator can perform no worse assuming the problem is well-posed (i.e. observability, etc.).

It is well known that state constraints may destabilize receding horizon controllers when the system is non-minimum phase, even when the constraints appear well-posed (c.f. (Zafiriou and Marchal

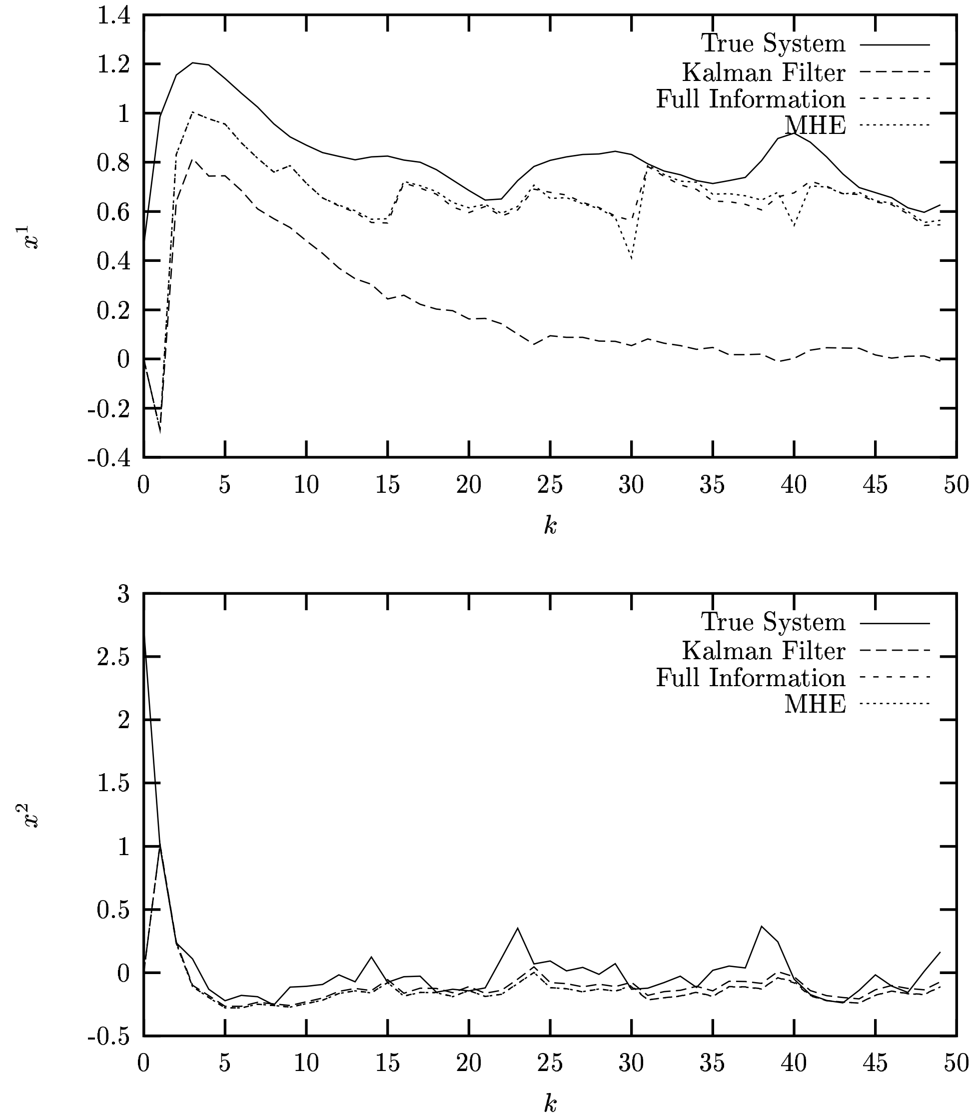


Figure 4.3: Comparison of estimators for Example 1

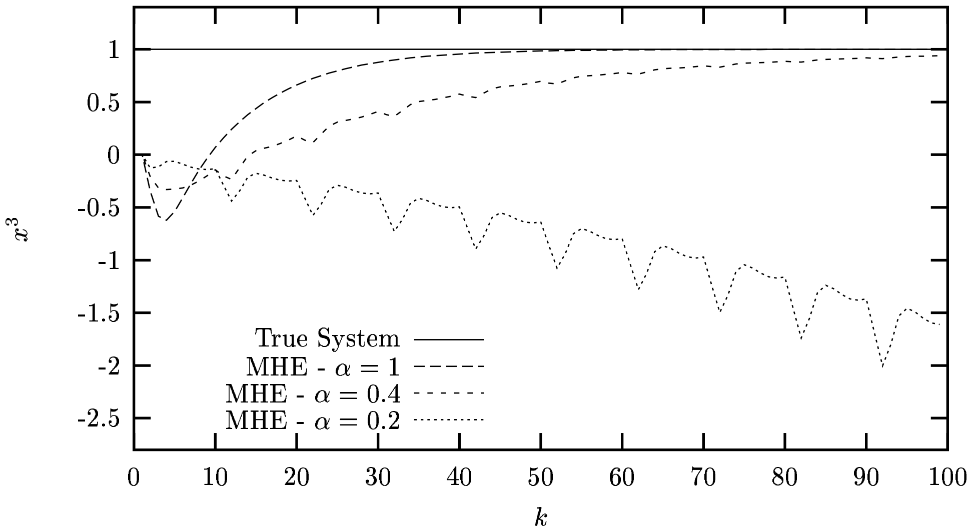


Figure 4.4: Example 2: comparison of steady-state MHE with different scalings of Π_∞.

1991)). The instability results from the controller inverting an unstable zero. Note this problem does not arise when the system is minimum phase. Rawlings and Muske (1993) demonstrated this form of instability results when the corresponding infinite horizon problem is infeasible. To alleviate these difficulties many researchers have proposed different strategies to ***soften*** the state constraints (c.f. (Ricker, Subrahmanian and Sim 1988, Genceli and Nikolaou 1993, de Oliveira and Biegler 1994, Zheng and Morari 1995, Scokaert and Rawlings 1996)). These strategies are also applicable to MHE *mutatis mutandis.*

One may reasonably expect similar forms of instability arise in MHE (one example is described below). We have already shown how instability may arise for non-minimum phase systems with an integrator if the initial penalty is improperly chosen. Instability arising from state constraints is a result of violating assumption **I** . We note assumption **I** is not a necessary condition for the stability of MHE, but it is a necessary condition for the full information estimator (otherwise, a solution does not exist in the limit as $T \to \infty$). From a practical point of view, we believe it suffices to consider the effect of constraints on the full information estimator only, because the stability of the full information estimator implies MHE is stable, and the problem we **desire** to solve is the full information problem.

Suppose assumption **I** is violated and the feasible region is non-empty (i.e. there exists sequence $\bar{x}_0$ and $\{\bar{w}_k\}_{k=0}^{\infty}$ satisfying the constraints (4.3)). Under these assumptions, it is unlikely one will be able to determine online that the problem is poorly posed, because the optimal cost ϕ_T^* is finite for all T. In regulation we do not have this limitation, because we can determine with the open-loop prediction whether a feasible infinite horizon solution exists using the theory of output admissible sets (Gilbert and Tan 1991). We do not, and cannot, have a similar theory in estimation, because it would violate causality and destroy the underlying framework of the estimation problem. Therefore, one must be careful with constraints in estimation. It is critical not to bias the estimate by adding unwarranted information. The constraints should be used only to add information to the estimation problem not available in the model equations. Examples include completing conservation equations (i.e. mass and energy are positive quantities) and modeling random variables sampled from truncated distributions.

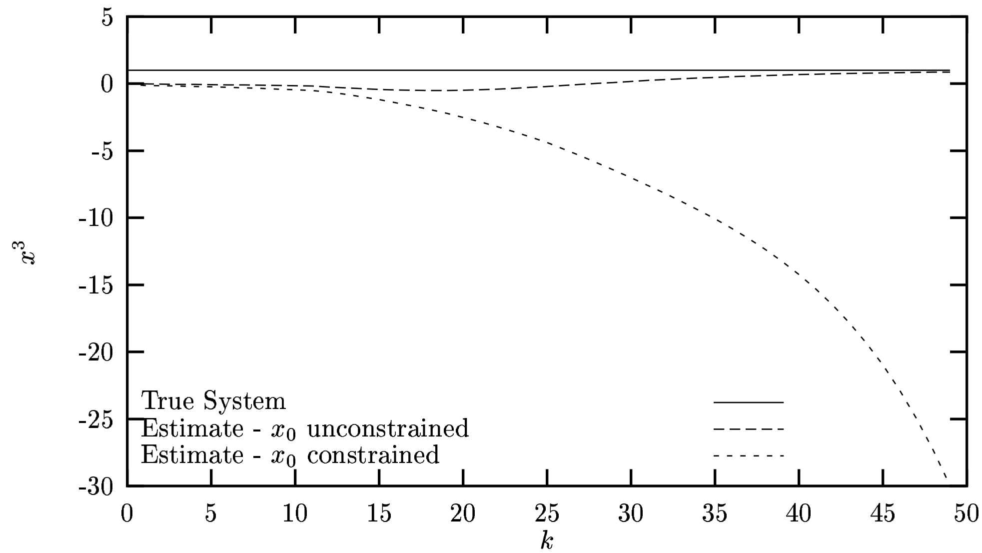


Figure 4.5: Example 3: instability due to constraints.

In the regulator, the state constraints are used to maintain the process in desirable operating regions. This situation does not arise in estimation, where we do not maintain but rather observe. Hence, we expect infeasibilities might arise in regulation due to process upsets, which is why many researchers have addressed constraint softening strategies for constrained regulation.

Consider again the system (4.10). We formulate the estimation problem with $Q = 1$, $R = 1$, $N = 5$, and the constraint $Cx_k \geq 0.1$. We restrict our attention to the full information estimator. Suppose we choose $\hat{x}_0 = 0$. This prior choice of the state is infeasible. However, the estimator can always choose a feasible x_0 at a cost $(x_0 - \hat{x}_0)'\Pi_0^{-1}(x_0 - \hat{x}_0)$. To introduce instability, we add the constraint $x_0 \leq 0$. To satisfy the output constraint, the estimator forces $x^{(3)} \to -\infty$. This response is precisely what we expect in receding horizon control when there is an infeasible output constraint. However, if we allow x_0 to have positive values, then the estimator is stable. Figure 4.5 shows the response of the full information estimator with the initial penalty $\Pi_0 = I$ and $x_0 \leq 0$. As expected, the estimator is unstable when we choose $x_0 \leq 0$. However, if we remove the constraint on x_0, then the estimator is stable as expected.

4.6.3 Example of MHE with a smoothing update.

Consider again the output constraint problem discussed in Example 4.6.2 where we estimate the state of the system (4.10) subject to the state constraint $Cx_k \geq 0.1$. Figure 4.6 shows a comparison of the full information estimator, MHE with a filter update, and MHE with a smoothing update with $\Pi_0 = I$. The kink at time $k = 5$ is due to the initialization of the smoothing updating strategy. Prior to $k = 5$, all of the estimators are equivalent. To understand why MHE with the smoothing update diverges from, and improves upon, the full information estimator and MHE with a filtering update, consider Figure 4.7. While $\hat{x}^{(3)}_{k|k-1}$ conveys no information about the step disturbance (between times $k = 0$ and $k = 27$, the estimate is negative), $\hat{x}^{(3)}_{j|k-1}$ for $j < (k-1)$ indicates a possible step disturbance. Unlike MHE with a filter update, this information is available to MHE with a smoothing update. Without constraints, all

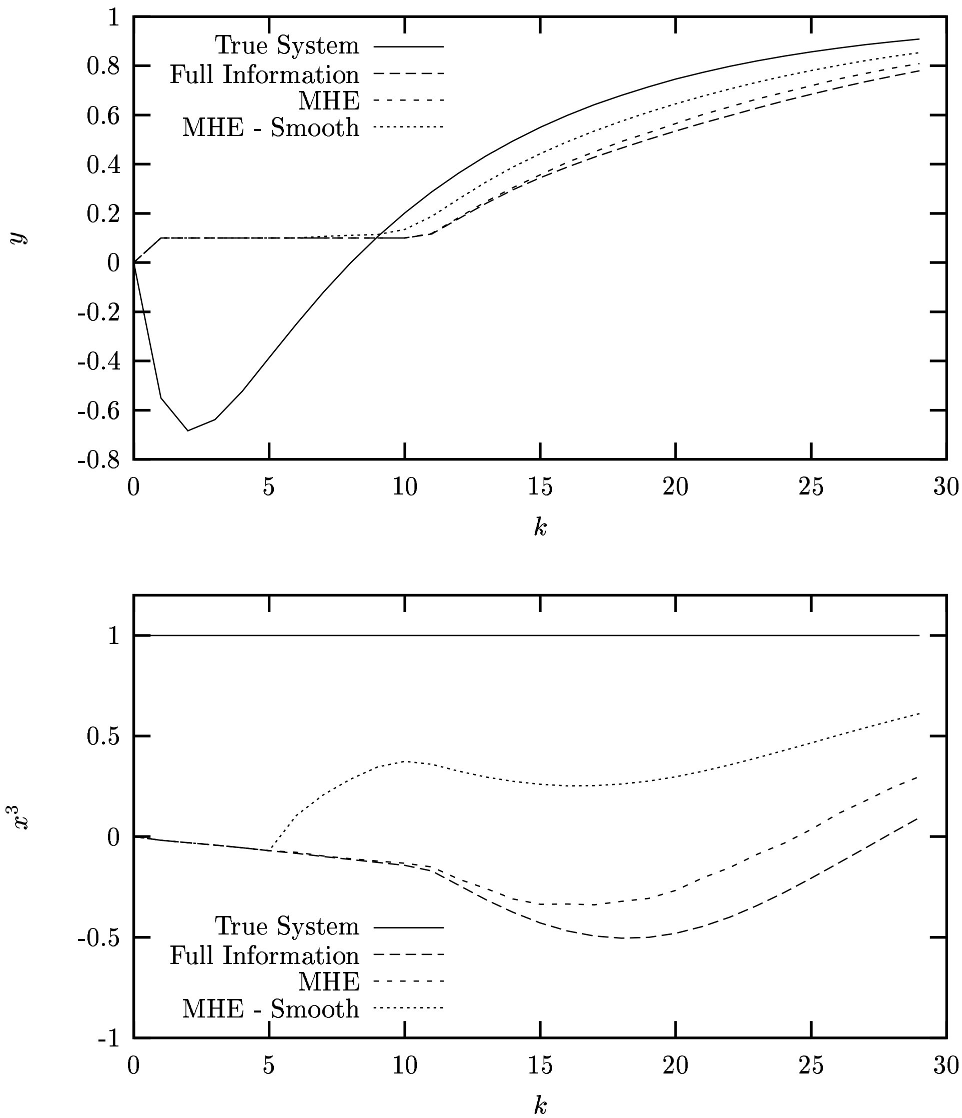


Figure 4.6: Example 4: comparison of MHE with a filtering and smoothing update.

three estimators are equivalent by construction.

One may view the constraint $Cx_k \geq 0.1$ as modeling error, because the true system does not obey the constraints. It is in these situations that MHE with a smoothing update may prove useful, because smoothing places more emphasis on recent information. As established by Findeisen (1997), a smoothing update improves the stability margins of unconstrained MHE. In a similar vein, the smoothing update in this example improves the *robustness* of the estimator subject to the spurious constraint by weighting less the past information. We do not make any general claims concerning the robustness of MHE with a smoothing update, only an observation that we believe is reasonable. Establishing and quantifying concrete benefits of smoothing updates are topics for future research.

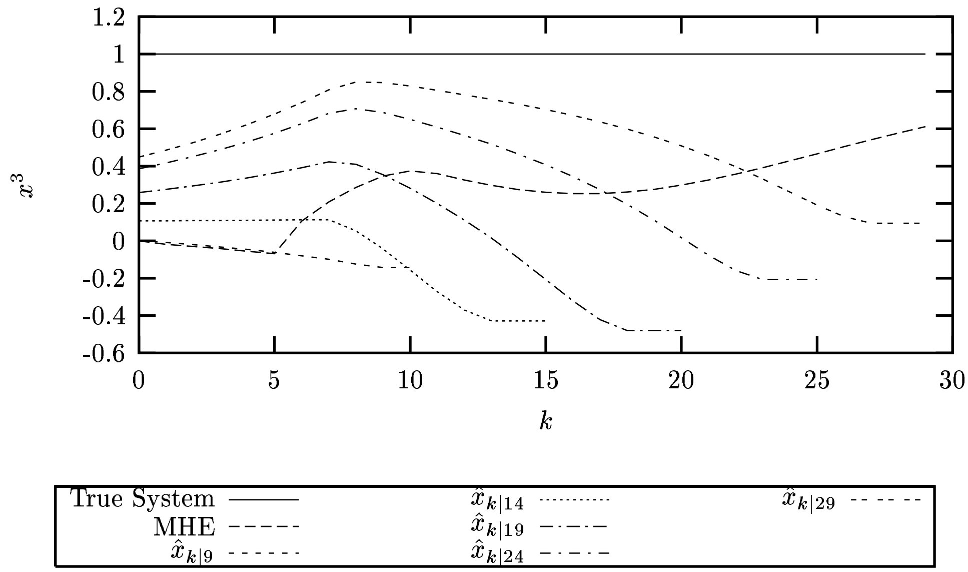


Figure 4.7: Example 4: comparison of MHE with smoothing update and the full information estimates of the state trajectory $\hat{x}_{\cdot|T}$ at different times T.

4.7 Conclusion

In this chapter we derived sufficient conditions for the stability of moving horizon estimation. Three separate formulations were presented. The key result of this work is that if the full information estimator is stable, then MHE is also stable provided one does not introduce extra bias with the prior information. To characterize this condition, we analyzed the estimation problem using forward dynamic programming using the techniques established in Chapter 3.

4.8 appendix

4.8.1 Proof of Lemma 4.4.2

Proof. Let x_0 denote the initial condition of the system (4.6) generating the output sequence $\{y_k\}$ (i.e. $y_k = CA^k x_0$) and $\epsilon > 0$. Using the state equation, we have the following expression for $\ell \leq N$:

$$\hat{x}_{T-N+\ell|T-1} = A^\ell \hat{x}_{T-N|T-1} + \sum_{j=0}^{\ell-1} A^{\ell-1-j} G \hat{w}_{T-N+j|T-1}.$$

Hence, we can derive the following bound for the estimation error:

$$\begin{aligned} \|\hat{x}_T - A^T x_0\| &= \|A^N(\hat{x}_{T-N|T-1} - A^{T-N}x_0) + \sum_{j=0}^{N-1} A^{N-1-j} G \hat{w}_{T-N+j|T-1}\|, \\ &\leq \|A\|^N \|\hat{x}_{T-N|T-1} - A^{T-N}x_0\| + \sum_{j=0}^{N-1} \|A\|^{N-1-j} \|G\| \|\hat{w}_{T-N+j|T-1}\|. \end{aligned} \quad (4.11)$$

Let $\hat{y}_{k|T-1} := C\hat{x}_{k|T-1}$. If we utilize the inverse triangle inequality, we obtain the following inequality

$$\begin{aligned}\sum_{j=0}^{N-1}\|\hat{v}_{T-N+j|T-1}\| &= \sum_{j=0}^{N-1}\|y_{T-N+j}-\hat{y}_{T-N+j|T-1}\|\\ &= \sum_{j=0}^{N-1}\|y_{T-N+j}-CA^j\hat{x}_{T-N|T-1}+CA^j\hat{x}_{T-N|T-1}-\hat{y}_{T-N+j|T-1}\|,\\ &\geq \sum_{j=0}^{N-1}\|y_{T-N+j}-CA^j\hat{x}_{T-N|T-1}\|-\sum_{j=0}^{N-1}\|CA^j\hat{x}_{T-N|T-1}-\hat{y}_{T-N+j|T-1}\|.\end{aligned}$$

If we rearrange the inequality, we obtain the new inequality

$$\sum_{j=0}^{N-1}\|\hat{v}_{T-N+j|T-1}\|+\sum_{j=0}^{N-1}\|CA^j\hat{x}_{T-N|T-1}-\hat{y}_{T-N+j|T-1}\|\geq\sum_{j=0}^{N-1}\|y_{T-N+j}-CA^j\hat{x}_{T-N|T-1}\|. \quad (4.12)$$

If we define the observability matrix

$$\mathcal{O} := \left[\begin{array}{cccc} C', & A'C', & \dots, & A^{N'}C' \end{array}\right]',$$

then

$$\sum_{k=0}^{N-1}\|CA^k x_0 - CA^k\hat{x}_0\|^2 = (x_0-\hat{x}_0)'\mathcal{O}'\mathcal{O}(x_0-\hat{x}_0).$$

We can derive the following bound

$$\sum_{j=0}^{N-1}\|CA^j x_0 - CA^j\hat{x}_0\|\geq\sqrt{\lambda_{\min}(\mathcal{O}'\mathcal{O})}\|x_0-\hat{x}_0\|.$$

The observability assumption guarantees $\lambda_{\min}(\mathcal{O}'\mathcal{O})>0$ for $N\geq n$. Hence, we have

$$\|\hat{x}_{T-N|T-1}-A^{T-N}x_0\|\leq\varphi\sum_{j=0}^{N-1}\|y_{T-N+j}-CA^j\hat{x}_{T-N|T-1}\|, \quad (4.13)$$

where

$$\varphi = \frac{1}{\sqrt{\lambda_{\min}(\mathcal{O}'\mathcal{O})}}.$$

So, if we substitute (4.13) and (4.12) into (4.11), we obtain

$$\begin{aligned}\|\hat{x}_{T|T-1}-A^T x_0\| \leq\ & \varphi\|A\|^N\left(\sum_{j=0}^{N-1}\|\hat{v}_{T-N+j|T-1}\|+\sum_{j=0}^{N-1}\|CA^j\hat{x}_{T-N|T-1}-\hat{y}_{T-N+j|T-1}\|\right)+\\ & \sum_{j=0}^{N-1}\|A\|^{N-1-j}\|G\|\|\hat{w}_{T-N+j|T-1}\|. \quad (4.14)\end{aligned}$$

Using the state equation, we obtain the following inequality for $j\leq N$:

$$\|CA^j\hat{x}_{T-N|T-1}-\hat{y}_{T-N+j|T-1}\|\leq\|C\|\sum_{\ell=0}^{j-1}\|A\|^{j-1-\ell}\|G\|\|\hat{w}_{T-N+\ell|T-1}\|. \quad (4.15)$$

Substituting (4.15) into (4.14) we derive the following bound on the estimation error:

$$
\begin{aligned}
\|\hat{x}_{T|T-1} - x_T\| &\leq \varphi\|A\|^N \left(\sum_{j=0}^{N-1} \|\hat{v}_{T-N+j|T-1}\| + \sum_{j=0}^{N-1} \|C\| \sum_{\ell=0}^{j-1} \|A\|^{j-1-\ell} \|G\| \|\hat{w}_{T-N+\ell|T-1}\| \right) + \\
&\quad \sum_{j=0}^{N-1} \|A\|^{N-1-j} \|G\| \|\hat{w}_{T-N+j|T-1}\|, \\
&\leq \varphi\|A\|^N \sum_{j=0}^{N-1} \Bigg(\|\hat{v}_{T-N+j|T-1}\| + \\
&\quad \|C\| \sum_{\ell=0}^{j-1} \|A\|^{j-1-\ell} \|G\| \|\hat{w}_{T-N+\ell|T-1}\| + \frac{\|A\|^{N-1-j}\|G\|}{\varphi\|A\|^N} \|\hat{w}_{T-N+j|T-1}\| \Bigg), \\
&\leq \varphi\|A\|^N \sum_{j=0}^{N-1} \left(1 + \|C\| \sum_{\ell=0}^{j-1} \|A\|^{j-1-\ell}\|G\| + \frac{\|A\|^{N-1-j}\|G\|}{\varphi\|A\|^N} \right) d^*,
\end{aligned}
$$

where

$$
d^* = \max_{j=T-N,\dots,T-1} \left\{ \|\hat{v}_{j|T-1}\|, \|\hat{w}_{j|T-1}\| \right\}.
$$

Without loss of generality, we assume A is not vacuous to generate the expressions for the last two inequalities. If we let $\eta = \min\{\lambda_{\min}(R^{-1}), \lambda_{\min}(Q^{-1})\}$ and choose ϱ such that

$$
\varrho \leq \eta \left(\frac{\epsilon}{\varphi\|A\|^N \sum_{j=0}^{N-1} \left(1 + \|C\| \sum_{\ell=0}^{j-1} \|A\|^{j-1-\ell}\|G\| + \frac{\|A\|^{N-1-j}\|G\|}{\varphi\|A\|^N} \right)} \right)^2,
$$

then $\|\hat{x}_T - A^T x_0\| \leq \epsilon$ when

$$
\sum_{k=N}^{T-1} \hat{v}'_{k|T-1} R^{-1} \hat{v}_{k|T-1} + \hat{w}'_{k|T-1} Q^{-1} \hat{w}_{k|T-1} \leq \varrho,
$$

and the Lemma follows as claimed. $\square$

4.8.2 Proof of Lemma 4.4.4

Before proving Lemma 4.4.4, we first establish the following lemma concerning general quadratic programs.

Lemma 4.8.1 *Let $\theta(z) = z'Qz$ where the matrix Q is symmetric positive definite and the sets Γ and Ω are closed and convex with $\Gamma \subseteq \Omega$. If a solution exists to the following quadratic programs $\theta(\hat{z}) = \min_{z\in\Omega} \theta(z)$, and $\theta(\bar{z}) = \min_{z\in\Gamma} \theta(z)$, then $\theta(\bar{z}) \geq \theta(\hat{z}) + \theta(\Delta z)$ where $\Delta z = \bar{z} - \hat{z}$.*

Proof. Substituting in for $\bar{z}$, we obtain

$$
\begin{aligned}
\theta(\bar{z}) &= \theta(\hat{z} + \Delta z) \\
&= \theta(\hat{z}) + \langle \nabla\theta(\hat{z}), \Delta z \rangle + \theta(\Delta z).
\end{aligned}
$$

Optimality implies $\langle \nabla\theta(\hat{x}), z - \hat{z}\rangle \geq 0$ for every $z \in \Omega$ [5]. This inequality implies $\theta(\bar{z}) \geq \theta(\hat{z}) + \theta(\Delta z)$ as claimed. $\square$

Proof. [Lemma 4.4.4] Without loss of generality, we take $\hat{x}_{T-N} = 0$. Consider an arbitrary $p \in \mathcal{R}_T$. Let

$$\left(\bar{x}_{T-N|T-1}, \{\bar{w}_{k|T-1}\}_{k=T-N}^{T-1}\right) = \arg \min_{z, \{w_k\}_{k=T-N}^{T-1}} \left\{\hat{\phi}_T(z, \{w_k\}) \ : \ x(N; z, \{w_j\}) = p\right\},$$

where the minimization is subject to the constraints (4.3). If

$$\Delta x_{T-N|T-1} := \bar{x}_{T-N|T-1} - \hat{x}_{T-N|T-1}, \qquad \Delta w_{k|T-1} := \bar{w}_{k|T-1} - \hat{w}_{k|T-1},$$

then, by Lemma 4.8.1, we have

$$\hat{\theta}_T(p) \geq \hat{\phi}_T^* + \hat{\phi}_T(\Delta x_{T-N|T-1}, \{\Delta w_{k|T-1}\}).$$

If we choose $p = \hat{x}_T$, then $\Delta x_{T-N|T-1} = 0$, $\Delta w_{k|T-1} = 0$, and

$$\hat{\phi}_T(\Delta x_{T-N|T-1}, \{\Delta w_{k|T-1}\}) = 0.$$

Let $\Delta p := p - \hat{x}_T$. We obtain, therefore, the following inequality

$$\begin{aligned} \hat{\phi}_T(\Delta x_{T-N|T-1}, \{\Delta w_{k|T-1}\}) &\geq \min_{\Delta z, \{\Delta w_j\}} \left\{\hat{\phi}(\Delta z, \{\Delta w_j\}) \ : \ x(N; \Delta z, \{\Delta w_j\}) = \Delta p\right\}, \\ &= (\Delta p)'\Pi_T^{-1}(\Delta p), \end{aligned}$$

and the lemma follows as claimed $\square$

4.8.3 Formulae for Smoothing Covariance

For notational simplicity, we make the following identities

$$\mathcal{O}_{N-2} := \begin{bmatrix} C' & A'C' & \cdots & A^{(N-2)'}C' \end{bmatrix}',$$

$$\mathcal{Y}_T^{N-2} := \begin{bmatrix} y_{T-N} \\ y_{T-N+1} \\ \cdots \\ y_{T-2} \end{bmatrix},$$

and

$$W_{N-2} =$$

$$\left[\begin{array}{ccccc} R & & & \cdots & \\ & CGQG'C' + R & CGQG'A'C' & \cdots & CGQG'A^{(N-3)'}C' \\ & CAGQG'C' & C(GQG' + AGQG'A')C' + R & & \\ \vdots & \vdots & & \ddots & \vdots \\ & CA^{(N-3)}GQG'C' & CA^{(N-3)}GQG'A'C' & \cdots & C(\sum_{k=0}^{N-3} A^k GQG' A^{k'})C' + R \end{array}\right].$$

[5] A proof by contradiction is immediate – assume there exists a z in Ω that violates the above condition and consider a convex combination between z and $\hat{z}$, which lies in Ω, and calculate cost; it decreases from $\theta(\hat{z})$ along the line, contradicting optimality. In other words, $-\nabla\theta(\hat{z}) \in T_\Omega(\hat{z})$ where $T_\Omega(\hat{z})$ denotes the normal cone to Ω at $\hat{z}$:

$$T_\Omega(\hat{z}) = \{z \ : \ \langle z^1 - \hat{z}, z\rangle \leq 0, \ \ \forall z^1 \in \Omega\}.$$

4.8.4 Proof of Lemma 4.5.2

Proof. In light of Lemma 4.4.4, it suffices to demonstrate

$$\Gamma_{T-N}(p) = (p - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(p - \hat{x}_{T-N}).$$

From Lemma 4.5.1, we have the following equality

$$\begin{aligned}
&(p - \hat{x}_{T-N|T-2})'\Pi_{T-N|T-2}^{-1}(p - \hat{x}_{T-N|T-2}) + \phi_{T-1}^{*} = \\
&\min_{\{w_k\}_{k=T-N}^{T-2}} \left\{ \sum_{k=T-N}^{T-2} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + (p - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(p - \hat{x}_{T-N}) \ : \right. \\
&\qquad\qquad \left. v_k = y_k - Cx(k - (T-N); p, \{w_j\}) \right\} + \hat{\phi}_{T-N}^{*}, \\
&= \min_{\{w_j\}_{T-N}^{T-2}} \left\{ \sum_{k=T-N}^{T-2} v_k' R^{-1} v_k + w_k' Q^{-1} w_k \ : \ v_k = y_k - Cx(k - (T-N); p, \{w_j\}) \right\} + \\
&(p - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(p - \hat{x}_{T-N}) + \hat{\phi}_{T-N}^{*}.
\end{aligned}$$

Let

$$D_{N-2}(p) := \min_{\{w_j\}_{T-N}^{T-2}} \left\{ \sum_{k=T-N}^{T-2} v_k' R^{-1} v_k + w_k' Q^{-1} w_k \ : \ v_k = y_k - Cx(k - (T-N); p, \{w_j\}) \right\}.$$

We may evaluate $D_{N-2}(p)$ using induction. Consider

$$D_1(p) = \min_{\{w_j\}_{T-3}^{T-2}} \left\{ \sum_{k=T-3}^{T-2} v_k' R^{-1} v_k + w_k' Q^{-1} w_k \ : \ v_k = y_k - Cx(k - (T-3); p, \{w_j\}) \right\}.$$

Evaluating the minimization analytically, we obtain

$$\begin{aligned}
D_1(p) &= \left(\begin{bmatrix} y_{T-3} \\ y_{T-2} \end{bmatrix} - \begin{bmatrix} C \\ CA \end{bmatrix} p \right)' \begin{bmatrix} R & 0 \\ 0 & R + CGQG'C' \end{bmatrix}^{-1} \left(\begin{bmatrix} y_{T-3} \\ y_{T-2} \end{bmatrix} - \begin{bmatrix} C \\ CA \end{bmatrix} p \right), \\
&= \left(\begin{bmatrix} y_{T-3} \\ y_{T-2} \end{bmatrix} - \begin{bmatrix} C \\ CA \end{bmatrix} p \right)' W_1^{-1} \left(\begin{bmatrix} y_{T-3} \\ y_{T-2} \end{bmatrix} - \begin{bmatrix} C \\ CA \end{bmatrix} p \right).
\end{aligned}$$

Now assume

$$D_{N-3}(p) = \left(\mathcal{Y}_T^{N-3} - \mathcal{O}_{N-3}p\right)' W_{N-3}^{-1} \left(\mathcal{Y}_T^{N-3} - \mathcal{O}_{N-3}p\right)$$

for the induction hypothesis and consider $D_{N-1}(p)$. A standard dynamic programming decomposition leads to the following reformulation

$$D_{N-2}(p) = \min_{w_{T-N}} \left\{ D_{N-3}(z) + v_{T-N}' R^{-1} v_{T-N} + w_{T-N}' Q^{-1} w_{T-N} \ : \ \begin{array}{c} v_{T-N} = Cp \\ z = x(1; z; w_{T-N}) \end{array} \right\}.$$

From the result concerning $D_1(p)$, we have the expression

$$D_{N-2}(p) = \left(\mathcal{Y}_T^{N-2} - \mathcal{O}_{N-2}p\right)' W_{N-2}^{-1} \left(\mathcal{Y}_T^{N-2} - \mathcal{O}_{N-2}p\right).$$

So,

$$\Gamma_{T-N}(p) = (p - \hat{x}_{T-N|T-2})'\Pi_{T-N|T-2}^{-1}(p - \hat{x}_{T-N|T-2}) - D_{N-2}(p) + (\hat{\phi}_{T-1}^{*} - \hat{\phi}_{T-N}^{*}).$$

By inspection

$$\Gamma_{T-N}(p) = (p - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(p - \hat{x}_{T-N}),$$

and the lemma follows as claimed. □

Chapter 5

Model Predictive Control

5.1 Introduction

The essence of model predictive control (MPC) is to optimize, over an open-loop time sequence of controls, the process response using a model to forecast the future process behavior over a prediction horizon. A conceptual picture of the prediction horizon is shown in Figure 5.1. Feedback is obtained by injecting the first control into the process and then resolving the optimization problem when new process measurements become available. The explicit use of a process model and prediction horizon are conceptually appealing to the practitioner and partially accounts for the popularity of MPC in industry.

While the concept of open-loop optimal feedback is relatively old (see for example (Dreyfus 1962)), MPC, or moving horizon control as it first was called, was first studied systematically in the automatic control community by Kleinman (1970), Thomas (1975), and Kwon and Pearson (1977, 1978). Whereas the initial interest in MPC by the automatic control community was limited to unconstrained linear time-varying systems, MPC was enthusiastically embraced by the process control community as the prevailing advanced control strategy to handle constrained multivariable systems. The application of MPC in process control was first documented by Richalet and coworker (1978) and Cutler and Ramaker (1980). The positive response of the process control community to MPC was due in part to these papers discussing successful industrial implementations rather than theoretical issues. A current survey of MPC is the process industries is given by Qin and Badgwell (1997, 1998), where they document over two thousand commercial applications of MPC.

Aside from the initial results for unconstrained linear time-varying systems, the first general stability result for MPC was established by Chen and Shaw (1982), who proved stability for unconstrained nonlinear systems using a terminal equality constraint. This result was further generalized by Mayne and Michalska (1990, 1991). Keerthi and Gilbert (1986) established the first stability result for constrained nonlinear systems again using a terminal equality constraint. This article is one of the key references in MPC theory, because it was the first to address constraints explicitly. Meadows and coworkers (1995) further demonstrated that MPC could stabilize systems that are feedback linearizable and also systems that cannot be stabilized by continuous feedback policies. The early formulations of MPC all relied on a terminal equality constraint for stability. Satisfying this constraint, however, is often computationally difficult, and optimization algorithms can only satisfy nonlinear equality constraints asymptotically. To relax the requirement of a terminal equality constraint, Michalska and Mayne (1993) established stability for constrained nonlinear systems using a terminal constraint set. They established also that optimality is not necessary for stability and proposed a suboptimal version of MPC. This article is the other key reference in MPC theory and is the foundation for most current MPC research. An alternative formulation of MPC using contraction constraints was proposed by Polak and Yang (1993*a*, 1993*b*, 1993). However, as discussed by Mayne (1997), feasibility issues may arise with contractive MPC.

In parallel to work done on constrained nonlinear stability, most process control researchers

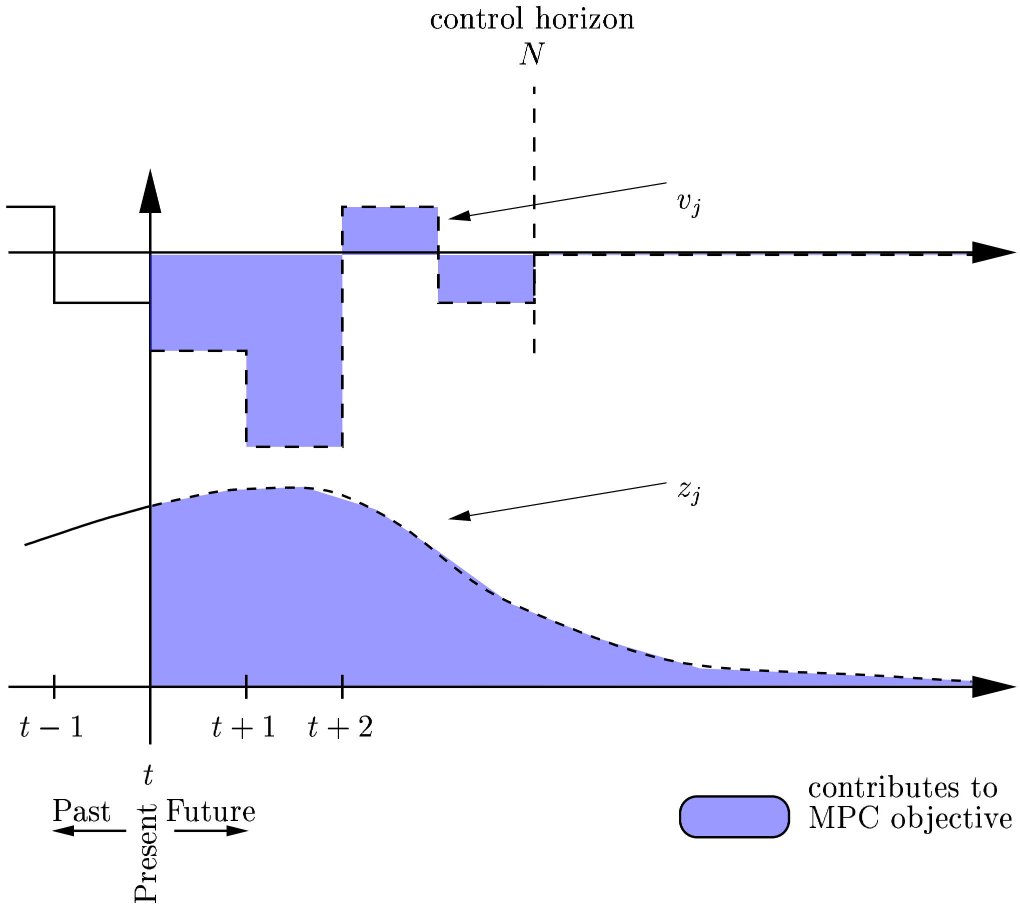


Figure 5.1: The prediction horizon in MPC

focused on constrained linear stability. Following the proposal of Bitmead, Gevers and Wertz (1990) for unconstrained linear systems, this work focused on modifying the terminal penalty for stability. The first proposals for constrained linear systems using a modified terminal penalty were made by Keerthi (1986) and Sznaier and Damborg (1987). These results were further developed (though independently) by Rawlings and Muske (1993). It is interesting to note a similar proposal was made earlier by Gauthier and Bornard (1983) for unconstrained linear systems. Further extensions were made by Chmielewski and Manousiouthakis (1996) and Scokaert and Rawlings (1998). Building on these results for constrained linear systems, many researchers have made analogous proposals for constrained nonlinear systems using a terminal constraint set combined with a modified terminal penalty. Examples include Parisini and Zoppoli (1995), Chisci, Lombardi and Mosca (1996), De Nicolao, Magni and Scattolini (1997), Chen and Allgöwer (1998), and Scokaert et al. (1999).

In addition to the research on MPC for constrained linear and nonlinear systems, there is a large amount of literature on MPC for unconstrained linear systems. Much of this literature focuses on generalized predictive control (GPC): MPC with transfer functions rather than state-space models. This literature is not reviewed, as it does not consider constraints. For a comprehensive review of the MPC literature, the reader is directed to the book by Camacho and Bordons (1998) and the survey papers by Garcia and coworkers (1989), Rawlings and coworkers (1994), Kwon (1996), Lee and Cooley (1997), Morari and Lee (1997), De Nicolao and coworkers (1998), Rawlings (1999), and Mayne and coworker (2000).

In this chapter we review the basic theory of nonlinear MPC. The details of (constrained) linear MPC are discussed in the subsequent chapters. The discussion is based on the review article by Mayne et al. (2000). We begin by defining MPC in Section 5.2. Our notation, definitions, and basic assumptions are introduced in Section 5.3. We then derive general stability results for MPC in Section 5.4 and show how these results encompass most MPC stability results discussed in the literature. We conclude by discussing computational issues and suboptimality in Section 5.5.

5.2 Problem Statement

Suppose the process is modeled by the following nonlinear discrete-time system

$$x_{k+1} = f(x_k, u_k), \tag{5.1}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $f(0,0) = 0$. Let $x(k; z, \{u_j\})$ denote the solution of the difference equation (5.1) at time k subject to the initial condition $x_0 = z$ at time 0 and control sequence $\{u_j\}_{j=0}^{k-1}$. We assume the physical limitations of the process (e.g. saturating valves) and the operating conditions (e.g. safety limits) require that the control and state sequences satisfy the following constraints

$$u_k \in \mathbb{U}, \qquad x_k \in \mathbb{X},$$

where the sets $\mathbb{U} \subseteq \mathbb{R}^m$ and $\mathbb{X} \subseteq \mathbb{R}^n$ are closed and contain the origin.

Consider the following open-loop optimal control problem

$$\mathcal{P}_N(z): \qquad V_N^*(z) = \min_{\{u_k\}_{k=0}^{N-1}} \{V_N(z, \{u_k\}) \ : \ \{u_k\} \in \mathcal{U}_N(z)\}$$

where the objective function is defined as

$$V_N(z, \{u_k\}) := \sum_{k=0}^{N-1} l(u_k, x_k) + F(x_N)$$

with $x_k := x(k; z, \{u_j\})$ and the constraint set is given by

$$\mathcal{U}_N(z) := \left\{ \{u_k\}_{k=0}^{N-1} \ : \ \begin{array}{l} u_k \in \mathbb{U}, \ \ k = 0, \ldots, N-1 \\ x(k; z, \{u_j\}) \in \mathbb{X}, \ \ k = 0, \ldots, N \\ x(N; z, \{u_j\}) \in \mathcal{X}_f \end{array} \right\}.$$

We assume the functions $l : \mathbb{U} \times \mathbb{X} \to \mathbb{R}$ and $F : \mathbb{X} \to \mathbb{R}$ and terminal constraint set $\mathcal{X}_f$ is closed and contains the origin. Let $\{u_k^*(z)\}_{k=0}^{N-1}$ denote the solution, assuming it exists, to $\mathcal{P}_N(z)$. We define model predictive control (MPC) as the feedback policy $\mu_N(\cdot) := u_0^*(\cdot)$. In particular, the first element $u_0^*(z)$ of the optimal sequence $\{u_k^*(z)\}_{k=0}^{N-1}$ is injected into the process. When a new measurement of the state z^+ becomes available, we solve $\mathcal{P}_N(z^+)$ and repeat the process. The terms "receding horizon" and "moving horizon" arise from the sliding prediction horizon (see Figure 5.1). MPC is called an open-loop feedback policy, because $\mathcal{P}_N(\cdot)$ is an open-loop optimal control problem; in particular, we optimize over fixed controls rather than feedback policies.

5.3 Notation, Definitions, and Basic Assumptions

The Cartesian product $\times_{k=1}^{N} \mathbb{A}$ of a set $\mathbb{A}$ is denoted by $\mathbb{A}^N$. We use the symbol $\|\cdot\|$ to denote any vector norm in $\mathbb{R}^n$ (where the dimension n follows from context). Let $\mathbb{R}_{\geq 0}$ denote the nonnegative real numbers and

$$f^{(k)}(\cdot) := \underbrace{f(f(\ldots f}_{k \text{ times}}(\cdot)\ldots)).$$

For $\epsilon > 0$, $N_\epsilon := \{x \;:\; \|x\| \leq \epsilon\}$.

Definition 5.3.1 *A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a* **K-function** *if it is continuous, strictly monotone increasing, $\alpha(x) > 0$ for $x \neq 0$, $\alpha(0) = 0$, and $\lim_{k\to\infty} \alpha(x) = \infty$.*

Fact 5.3.2 *Suppose $\alpha(\cdot)$ is a K-function. Then, the function $\alpha(\cdot)$ and its inverse $\alpha^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ are continuous (Royden 1988). Furthermore, $\alpha^{-1}(\cdot)$ is a K-function.*

Definition 5.3.3 *A system is* **constrained stabilizable** *if, for all $z \in \mathbb{X}$, there exists an input sequence $\{\bar{u}_k\}_{k=0}^{\infty} \in \mathcal{U}_\infty(z)$ and a K-function $\varphi(\cdot)$ such that*

$$\sum_{k=0}^{\infty} \|(\bar{u}_k, \bar{x}_k)\| \leq \varphi(\|z\|),$$

where $\bar{x}_k := x(k; z, \{\bar{u}_j\})$.

In order to guarantee the problem $\mathcal{P}_N(\cdot)$ is well-posed, we assume that the functions $f(\cdot)$, $l(\cdot)$ and $F(\cdot)$ satisfy the following conditions.

A0 The function $f(\cdot)$ is continuous.

A1 The functions $l(\cdot)$ and $F(\cdot)$ are lower semi-continuous.

A2 There exist K-functions $\eta(\cdot)$ and $\gamma(\cdot)$ such that

$$\begin{aligned} \eta(\|(u,x)\|) \leq l(u,x) \leq \gamma(\|(u,x)\|) \\ \eta(\|(x)\|) \leq F(x) \leq \gamma(\|(x)\|) \end{aligned}$$

for all $u \in \mathbb{U}$ and $x \in \mathbb{X}$.

Definition 5.3.4 *Consider the system*

$$x_{k+1} = f(x_k), \tag{5.2}$$

and let $x(k; z)$ denote the solution of the difference equation (5.2) at time k subject to the initial condition $x_0 = z$ at time 0. The system (5.2) is **stable** *if, for all $\epsilon > 0$, there exists a number $\delta > 0$ and a positive integer $\tilde{T}$ such that if $x_0 \in \mathbb{X} \cap N_\delta$, then $x(k; x_0) \in N_\epsilon$ for all $k \geq \tilde{T}$.*

To prove a solution exists to $\mathcal{P}_N(x)$, we need to characterize the set of initial states $\mathcal{X}_N$ for which the set $\mathcal{X}_f$ is reachable from in time N. Let

$$\mathcal{X}_N := \left\{x \;:\; \exists \{u_k\}_{k=0}^{N-1} \in \mathcal{U}_N(x)\right\}.$$

If $\mathcal{X}_f$ is not reachable from any state in time N, then $\mathcal{X}_N = \emptyset$. Likewise, if $\mathcal{X}_f$ is reachable from all states in time N, then $\mathcal{X}_N = \mathbb{R}^n$. If the system (5.1) is linear (i.e. $x_{k+1} = Ax_k + Bu_k$), $N \geq n$, the control is unconstrained (i.e. $\mathbb{U} = \mathbb{R}^m$), and rows of the matrix

$$\left[B,\; AB,\; A^2B, \ldots,\; A^{N-1}B\right]$$

are linearly independent (i.e. the pair (A, B) is controllable), then, for all $\mathcal{X}_f \subseteq \mathbb{R}^n$, $\mathcal{X}_N = \mathbb{R}^n$.

5.4 Stability

Optimality does not imply stability (Kalman 1960*a*). Additional measure are necessary, therefore, to guarantee stability. One typically guarantees stability by forecasting over an infinite horizon (solve the problem $\mathcal{P}_\infty(\cdot)$). Both the linear quadratic regulator (LQR) and $\mathcal{H}_\infty$ control implicitly employ an infinite horizon. However, for nonlinear systems, it is often impossible to forecast over an infinite horizon: to do so, one needs to evaluate an infinite summation. Furthermore, when a system possesses unstable dynamics, it is necessary to optimize over an infinite sequence of controls. This requirement is computationally impossible, unless the infinite sequence of controls can be represented with a finite basis. For linear systems, such a basis exists. However, for nonlinear systems, no universal basis exists unless one employs the terminal equality constraint $\mathcal{X}_f = \{0\}$.

The standard strategy to formulate a stabilizing MPC feedback policy (contractive MPC being the main exception) is to employ, either explicitly or implicitly, a local feedback policy $\kappa_f(\cdot)$ in conjunction with the terminal penalty $F(\cdot)$ and terminal constraint set $\mathcal{X}_f$. Typically, $\kappa_f(\cdot)$ is a stable linear feedback policy (i.e. $\kappa_f(x) = Kx$) for the linearized system $x_{k+1} = Ax_k + Bu_k$, where

$$A := \left.\frac{\partial f(x,0)}{\partial x}\right|_{x=0}, \qquad B := \left.\frac{\partial f(0,u)}{\partial u}\right|_{u=0}.$$

The terminal penalty $F(\cdot)$ is typically chosen as the local (quadratic) Lyapunov function for the system $x_{k+1} = (A+BK)x_k$. If the pair (A, B) is controllable and the function $f(\cdot)$ is sufficiently smooth, then there exists an $\alpha > 0$ and a positive definite matrix P such that the level set $\mathcal{L}_\alpha = \{x \,:\, x^T P x \leq \alpha\}$ is positive invariant for the system $x_{k+1} = f(x_k, \kappa_f(x_k))$ (Sontag 1990). In other words, the local feedback policy $\kappa_f(\cdot)$ will stabilize the system (5.1) for all $x_0 \in \mathcal{L}_\alpha$. If the terminal penalty $F(\cdot)$ is a local Lyapunov function for the system $x_{k+1} = f(x_k, \kappa_f(x_k))$ and the terminal constraint set $\mathcal{X}_f$, typically a level set of the terminal penalty $F(\cdot)$, is chosen such that it is positive invariant and $\mathcal{X}_f \subset \mathbb{X} \cap \mathbb{X}_\kappa$, where $\mathbb{X}_\kappa = \{x \,:\, \kappa_f(x) \in \mathbb{U}\}$, then the set $\mathcal{X}_f$ is output admissible (i.e. $\kappa_f(\cdot)$ satisfies the constraints $\mathbb{U}$ and $\mathbb{X}$ for all $x \in \mathcal{X}_f$). The terminal penalty and the terminal constraint set, implicitly through the local feedback policy $\kappa_f(\cdot)$, may be used then to approximate the tail of an infinite prediction horizon; i.e., for all $x \in \mathcal{X}_f$,

$$F(z) \approx \sum_{k=0}^{\infty} l(x_k, \kappa_f(x_k))$$

where x_k denotes the solution of the system $x_{k+1} = f(x_k, \kappa_f(x_k))$ at time k subject to the initial condition $x_0 = z$ at time 0. This construction allows one to approximate an infinite-horizon controller with a finite-horizon controller. So long as the approximation errors are small, infinite-horizon properties such as closed-loop stability result.

To guarantee stability, we require that the terminal penalty $F(\cdot)$, local feedback policy $\kappa_f(\cdot)$, and terminal constraint set $\mathcal{X}_f$ satisfy the following conditions.

C0 $\kappa_f(x) \in \mathbb{U}$ for all $x \in \mathcal{X}_f$. (output admissibility)

C1 $f(x, \kappa_f(x)) \in \mathcal{X}_f$ for all $x \in \mathcal{X}_f$. (positive invariance)

C2 $F(x) \geq F(f(x, \kappa_f(x)) + l(\kappa_f(x), x)$ for all $x \in \mathcal{X}_f$. ($F(\cdot)$ is a local Lyapunov equation)

As we demonstrate at the end of this section, by appropriately defining $\kappa_f(\cdot)$, $F(\cdot)$, and $\mathcal{X}_f$, most (stable) formulations of MPC satisfy conditions **C0**–**C2**. Consequently, we can generalize many MPC stability results by proving stability for conditions **C0**–**C2**.

Proposition 5.4.1 *If assumptions* **A0-A2** *hold, then a solution exists to* $\mathcal{P}_N(x)$ *for all* $x \in \mathcal{X}_N$.

Proof. By the assumption $x \in \mathcal{X}_N$, there exists $\{\bar{u}_k\} \in \mathbb{U}_N(x)$. Let $V_N^1(x) = V_N(x, \{\bar{u}_j\})$ and consider the set

$$\Lambda = \left\{\{u_k\}_{k=0}^{N-1} \ : \ \{u_k\} \in \mathcal{U}_N(z), \ \ V_N(x, \{u_k\}) \leq V_N^1\right\}.$$

By assumption **A0** and **A1**, the objective function $V_N(x, \cdot)$ is lower semi-continuous. The set

$$\Omega = \left\{\{u_k\}_{k=0}^{N-1} \ : \ V_N(x, \{u_k\}) \in [0, V_N^1]\right\}$$

is closed, because the inverse image of a closed set under a lower semi-continuous function is closed (Berge 1963). The set $\mathcal{U}_N(x)$ is closed, because the function $f(\cdot)$ is continuous and the sets $\mathbb{U}$, $\mathbb{X}$, and $\mathcal{X}_f$ are closed. The set Λ is closed, because it is the intersection of the closed sets Ω and $\mathcal{U}_N(z)$. By assumption **A2**, there exists a K-function $\eta(\cdot)$ such that $l(u, x) \geq \eta(\|u, x\|)$ and $F(x) \geq \eta(\|x\|)$. These inequalities imply the set F is also bounded, because $F(x_N) \leq V_N^1$ and $l(u_k, x_k) \leq V_N^1$ for $k = 0, \ldots, T-1$. The set Λ is bounded, because $\Lambda \subseteq \Omega$. Hence, the set Λ is compact. Existence of a solution follows from the Weierstrass Maximum Theorem. $\square$

Let

$$x_{k+1} = G(x_k) := f(x_k, \mu_N(x_k))$$

denote the closed-loop dynamics of the system (5.1) subject to the MPC feedback policy $\mu_N(\cdot)$. The following proposition establishes stability by demonstrating that the cost function $V_N^*(\cdot)$ is a Lyapunov function for the system $x_{k+1} = G(x_k)$.

Proposition 5.4.2 *Suppose $F(\cdot)$, $\kappa_f(\cdot)$, and $\mathcal{X}_f$ satisfy conditions* **C0**-**C2**, *assumptions* **A0**-**A2** *hold, and the system (5.1) is constrained stabilizable, then, for all $N \geq 1$, the system*

$$x_{k+1} = G(x_k)$$

is stable and $G^{(k)}(z) \to 0$ as $k \to \infty$ for all $z \in \mathcal{X}_N$.

Proof. Existence of a solution to $\mathcal{P}_N(\cdot)$ is established in Proposition 5.4.1. We first demonstrate convergence. Let $z \in \mathcal{X}_N$. We know $\{u_k^*(z\} \in \mathcal{U}_N(z)$ implies

$$\{u_1^*(z), u_2^*(z), \ldots, u_{N-1}^*(z)\} \in \mathcal{U}_{N-1}(G(z)).$$

Let $x_N^*(z) = x(N; z, \{u_j^*(z)\})$. Existence implies $x_N^*(z) \in \mathcal{X}_f$ and condition **C1** implies

$$f(x_N^*(z), \kappa_f(z)) \in \mathcal{X}_f.$$

Hence,

$$\{u_1^*(z), u_2^*(z), \ldots, u_{N-1}^*(z), \kappa_f(x_N(z))\} \in \mathcal{U}_N(G(z)).$$

Optimality implies the following inequality

$$V_N^*(z) - V_N^*(G(z)) \geq l(z, u_0^*(z)) + F(x_N^*(z)) - (l(\kappa_f(x_N(z)), x_N^*(z)) + F(f(x_N^*(z), \kappa_f(x_N^*(z))))) .$$

Condition **C2** implies

$$V_N^*(z) \geq l(\kappa_f(z), z) + V_N^*(G(z)). \tag{5.3}$$

As $V_N^*(\cdot)$ is bounded below (by assumption **A2**), the inequality (5.3) implies that the sequence $V_N^*(G^{(k)}(z))$ converges. Therefore,

$$l(G^{(k)}(z), \mu(G^{(k)}(z)) \to 0$$

as $k \to \infty$, and, by assumption **A2**, $G^{(k)}(z) \to 0$ as $k \to 0$.

To prove stability, let $\epsilon > 0$ and $\tilde{T} = 0$. Optimality and assumption **A2** imply

$$(N+1)\eta(\varphi(z)) \geq V_N^*(z) \geq \eta(\|z\|).$$

Likewise, the above inequality and inequality (5.3) imply

$$(N+1)\eta(\varphi(z)) \geq V_N^*(G(z)) \geq \eta(\|G(z)\|).$$

Hence, we have

$$(N+1)\eta(\varphi(z)) \geq \eta(\|G^{(k)}(z)\|)$$

for all $k \geq 0$. If we choose

$$\delta < \varphi^{-1}\left(\gamma^{-1}\left(\frac{\eta(\epsilon)}{N+1}\right)\right),$$

where the existence of the K-functions $\gamma^{-1}(\cdot)$ and $\varphi^{-1}(\cdot)$ follows from Fact 5.3.2, then the proposition follows as claimed. □

5.4.1 Case Examples

In this section, we apply the results of Proposition 5.4.2 to some MPC formulations discussed in the literature.

Infinite-Horizon MPC

Consider the infinite-horizon open-loop optimal control problem

$$\mathcal{P}_\infty(z): \qquad V_\infty^*(z) = \min_{\{u\}_{k=0}^\infty} \left\{ \sum_{k=0}^{\infty} l(u_k, x_k) \ : \ \{u_k\} \in \mathcal{U}_\infty(z), \ \ x_k := x(k; z, \{u_j\}) \right\}$$

and the associated infinite-horizon MPC feedback law $\mu_\infty(\cdot) := u_0^\infty(z)$, where $\{u_k^\infty(z)\}_{k=0}^\infty$ denotes the solution to $\mathcal{P}_\infty(z)$. Infinite-horizon MPC was first discussed by Keerthi and Gilbert (1988). This result has only theoretical significance, because $\mathcal{P}_\infty(\cdot)$ is, in general, impossible to solve. However, using the principle of optimality, we can recover infinite-horizon MPC from finite-horizon MPC if we choose the terminal penalty $F(\cdot) = V_\infty(\cdot)$ and $\mathcal{X}_f = \mathbb{X}$. This result is important when we consider constrained linear MPC.

Proposition 5.4.3 *Suppose assumptions* **A0-A2** *are true and the system (5.1) is constrained stabilizable, then $\mathcal{P}_\infty(z)$ has a solution for all $z \in \mathbb{X}$.*

Proof. The stated assumptions satisfy the infinite horizon existence result (Theorem 2) of Keerthi and Gilbert (1985). □

Remark 5.4.4 *Appendix A establishes existence and uniqueness when the system (5.1) is linear, the objectives are quadratic, and the sets $\mathbb{U}$ and $\mathbb{X}$ are convex.*

Remark 5.4.5 *Proposition 5.4.3 guarantees $F(\cdot)$ is defined on $\mathbb{X}$.*

Lemma 5.4.6 *Suppose assumptions* **A0-A2** *are true and the system (5.1) is constrained stabilizable. Then, $V_N^*(x) = V_\infty^*(x)$ for all $x \in \mathbb{X}$. Furthermore, $\{u_k^\infty(z)\}_{k=0}^{N-1}$ is a solution to $\mathcal{P}_N(z)$.*

Proof. As $\mathcal{X}_f = \mathbb{X}$, constrained stabilizability guarantees $\mathcal{X}_N = \mathbb{X}$. The remaining steps of the proof involve decomposing the problem $\mathcal{P}_\infty$ using dynamic programming and the principle of optimality (c.f. Bertsekas (1995*a*, 1995*b*)). □

Lemma 5.4.7 *Suppose assumptions* **A0-A2** *are true and the system (5.1) is constrained stabilizable. If* $F(\cdot) = V_\infty(\cdot)$*,* $\mathcal{X}_f = \mathbb{X}$*, and* $\kappa_f = \mu_\infty(\cdot) := u_0^\infty(\cdot)$*, then conditions* **C0-C2** *are satisfied.*

Proof. Proposition 5.4.3 implies $u_0^\infty(z) \in \mathbb{U}$ for all $x \in \mathbb{X}$, and, therefore, condition **C0** is satisfied. $\{u_k^\infty(z)\}_{k=0}^\infty \in \mathcal{U}_\infty(z)$ guarantees condition **C1** is satisfied. Using the principle of optimality, we know that the value function $V_\infty(\cdot)$ satisfies the following discrete-time Hamilton-Jacobi-Bellman equation

$$\begin{aligned} F(z) &= \min_u \{l(u,z) + F(f(z,u))\,;\, u \in \mathbb{U},\ \ f(z,u) \in \mathbb{X}\}, \\ &= l(u_0^\infty(z), z) + F(f(z, u_0^\infty(z))), \end{aligned}$$

for all $x \in \mathbb{X}$. Consequently, $F(\cdot)$ satisfies **C2** by equality. $\square$

Corollary 5.4.8 *Suppose assumptions* **A0-A2** *are true and the system (5.1) is constrained stabilizable. Then, infinite-horizon MPC is a stable control policy for the system (5.1) and* $G^{(k)}(z) \to 0$ *as* $k \to \infty$ *for all* $z \in \mathbb{X}$*..*

Corollary 5.4.9 *Suppose assumptions* **A0-A2** *are true, the system (5.1) is constrained stabilizable, and* $N \geq 1$*. If* $F(\cdot) = V_\infty(\cdot)$*,* $\mathcal{X}_f = \mathbb{X}$*, and* $\kappa_f = \mu_\infty(\cdot) := u_0^\infty(\cdot)$*, then finite-horizon MPC is a stable control policy for the system (5.1) and* $G^{(k)}(z) \to 0$ *as* $k \to \infty$ *for all* $z \in \mathbb{X}$*.*

Remark 5.4.10 *For constrained linear systems, it is possible to provide an analytic expression for* $V_\infty(\cdot)$ *in a neighborhood of the origin. The details are discussed in Chapter 6.*

Terminal-State MPC

One strategy to forecast over an infinite horizon is to employ a terminal equality constraint $\mathcal{X}_f = \{0\}$. In this case, the infinite-time behavior of the process is known; if we choose $\{u_k^*(\cdot)\}_{k=N}^\infty = 0$, then $x_N = 0$ and $f(0,0) = 0$ imply $x_k = 0$ for $k \geq N$. We may embed this strategy in our framework if we define $F(\cdot) = 0$, $\kappa_f(\cdot) = 0$, and $\mathcal{X}_f = \{0\}$.

Lemma 5.4.11 *If* $F(\cdot) = 0$*,* $\kappa_f(\cdot) = 0$*, and* $\mathcal{X}_f = \{0\}$*, then conditions* **C0-C2** *are satisfied.*

Proof. As $0 \in \mathbb{U}$, $0 \in \mathbb{X}$, $f(0,0) = 0$, and $l(0,0) = 0$, conditions **C0-C1** are satisfied trivially. $\square$

Corollary 5.4.12 *Suppose assumptions* **A0-A2** *are true, the system (5.1) is constrained stabilizable, and* $N \geq 1$*. Then, terminal-state MPC is a stable control policy for the system (5.1) and* $G^{(k)}(z) \to 0$ *as* $k \to \infty$ *for all* $z \in \mathcal{X}_N$*.*

Quasi-Infinite-Horizon MPC

In addition to their stability properties, infinite-horizon control laws have the property that the open-loop predictions are identical to the closed-loop response in *nominal* application. The effect of the controller's tuning parameters (parameterizing the stage costs $l(\cdot)$) are, therefore, more intuitive for infinite-horizon formulations than for finite-horizon formulations, where the open-loop predictions do not necessarily match the closed-loop response. As the goal of the prediction horizon is to anticipate the closed-loop response, one should design the controller such that the finite-horizon predictions approximate the infinite-horizon predictions.

One can approximate infinite-horizon MPC using the linearized dynamics of the system (5.1) (c.f. (Parisini and Zoppoli 1995, Chen and Allgöwer 1998)). We refer to this strategy as quasi-infinite-horizon MPC. Most recent proposals for nonlinear MPC are based on this idea. Suppose

$$l(u,x) = \frac{1}{2}\left(u^T R u + x^T Q x\right),$$

where the matrices R and Q are positive definite, and let

$$A := \left.\frac{\partial f(x,0)}{\partial x}\right|_{x=0}, \qquad B := \left.\frac{\partial f(0,u)}{\partial u}\right|_{u=0}.$$

We assume the pair (A,B) is controllable. If we consider the unconstrained linear system $x_{k+1} = Ax_k + Bu_k$, then

$$z^T P z = \min_{\{u\}_{k=0}^{\infty}} \left\{ \sum_{k=0}^{\infty} u_k^T R u_k + x_k^T R x_k \ : \ x_0 = z, \ \ x_{k+1} = Ax_k + Bu_k \right\},$$

where the matrix P is the unique positive definite solution of the following algebraic Riccati equation

$$P = Q + A^T \left(P - PB(R + B^T PB)^{-1} B^T P\right) A.$$

If the state z is sufficiently close to the origin and the function $f(\cdot)$ is sufficiently smooth, then the linearized dynamics accurately describe the behavior of the nonlinear system $x_{k+1} = f(x_k, u_k)$, and we have

$$z^T P z \approx V_\infty(z)$$

assuming the nonlinear problem is unconstrained: $\mathbb{U} = \mathbb{R}^m$ and $\mathbb{X} = \mathbb{R}^n$. Furthermore, the optimal (unconstrained) linear solution is close to the optimal (unconstrained) nonlinear solution; i.e.

$$\begin{aligned} u_0(z) &= (R + B^T PB)^{-1} B^T PAz, \\ &:= Kz, \\ &\approx u_0^\infty(z). \end{aligned}$$

Using these elementary results, we can approximate *constrained* infinite-horizon MPC with finite-horizon MPC. If we choose $F(x) = x^T Px$, $\kappa_f(x) = Kx$, and

$$\mathcal{X}_f = \{x \ : \ x^T Px \leq \alpha\},$$

where $\alpha > 0$ chosen such that

$$\max_{x \in \mathcal{X}_f} \{F(f(x, \kappa_f(x)) - F(x) + (1/2)x^T(Q + K^T RK)x\} \leq 0 \tag{5.4}$$

and $\mathcal{X}_f \subset \mathbb{X}_K := \{x \ : \ x \in \mathbb{X}, \ Kx \in \mathbb{U}\}$, then conditions **C0-C2** are satisfied. The existence of such a number $\alpha > 0$ is guaranteed provided $f(\cdot)$ is twice continuously differentiable, the pair (A,B) is controllable, and there exist numbers $\epsilon_1 > 0$ and $\epsilon_2 > 0$ such that $N_{\epsilon_1} \subset \mathbb{U}$ and $N_{\epsilon_2} \subset \mathbb{X}$ (Scokaert et al. 1999).

Lemma 5.4.13 *If $F(x) = x^T Px$, $\kappa_f(x) = K(x)$, and $\mathcal{X}_f = \{x : x^T Px \leq \alpha\}$ where $\alpha > 0$ satisfies the inequality (5.4) and the inclusion $\mathcal{X}_f \subset \mathbb{X}_K$, then conditions* **C0-C2** *are satisfied.*

Proof. The conditions **C0-C2** are satisfied by construction: $x \in \mathcal{X}_f$ implies $\kappa_f x \in \mathbb{U}$ and

$$\begin{aligned} \alpha &\geq x^T Px \\ &\geq f(x, \kappa_f(x))^T Pf(x, \kappa_f(x)) + (1/4)x^T(Q + K^T RK)x \\ &\geq f(x, \kappa_f(x))^T Pf(x, \kappa_f(x)). \end{aligned}$$

Consequently, conditions **C0** and **C1** are satisfied. If we substitute in for $\kappa_f(\cdot)$, then

$$l(\kappa_f(x), x) = \frac{1}{2} x^T \left(Q + K^T RK\right) x$$

and **C2** is satisfied by inequality (5.4). □

Corollary 5.4.14 *Suppose assumptions* **A0-A2** *are true, system (5.1) is constrained stabilizable,* $f(\cdot)$ *is twice continuously differentiable, the pair* (A, B) *is controllable, and* $N \geq 1$. *Then, quasi-infinite-horizon MPC is a stable control policy for the system (5.1) and* $G^{(k)}(z) \to 0$ *as* $k \to \infty$ *for all* $z \in \mathcal{X}_N$.

Remark 5.4.15 *When implementing nonlinear MPC, one typically ignores the terminal constraint set in the optimization problem. Instead, one choose the horizon length sufficiently large such that* $x_N \in \mathcal{X}_f$. *If we assume the set* $\mathcal{X}$ *is compact, then it is straightforward to show the existence of an integer* N *such that* $x(N; z, \{u_j^*(z)\}) \in \mathcal{X}_f$ *for all* $z \in \mathcal{X}_N$. *This strategy was employed by Alamir and Bornard (1995), Parisini and Zoppoli (1995), and Jadbabaie, Yu and Hauser (1999) for unconstrained nonlinear systems.*

5.5 Computational Issues and Suboptimality

If the system 5.1 is linear, the cost functions quadratic, and the sets $\mathbb{U}$ and $\mathcal{X}$ are polyhedral convex sets, then the optimal control problem $\mathcal{P}_N(\cdot)$ is quadratic program. Efficient software exists for the solution of quadratic programs. We can improve the efficiency of the optimization algorithms by exploiting the sparse structure of $\mathcal{P}_N(\cdot)$. Chapter 8 discusses one such approach. If the system (5.1) is nonlinear, then the optimal control problem $\mathcal{P}_N(\cdot)$ is a nonlinear program. Many algorithms exist for solving nonlinear programs (c.f. Nocedal and Wright (1999)). However, the problem is nonconvex, and optimization may not yield global solutions. Global optimization is an option; however, unless global information such as Lipschitz constants are available, the algorithm needs to sample a dense subset of the decision space (Stephens and Baritompa 1998). The algorithm requires this information to verify whether a particular solution is global. As global information for the problem $\mathcal{P}_N(\cdot)$ is rarely available, we need to guarantee stability in the absence of optimality. Michalska and Mayne (1993) and Mayne (1995*a*) demonstrate that optimality is not necessary for stability; rather, feasibility suffices. These ideas were developed further by Scokaert et al. (1999).

Finding a feasible element is itself a nonconvex optimization problem, though it is easy to verify whether a given element is feasible. One, therefore, can circumvent the computational difficulties of global optimization by implementing a *suboptimal* version of MPC.

Algorithm: Suboptimal MPC

Data $\alpha \in (0, 1]$.

Step 0 Given state x_0 at time $j = 0$, find a control sequence $\{u_k(x_0)\}_{k=0}^{N-1} \in \mathcal{U}_N(x_0)$. Let $V_N'(x_0) = V_N(x_0, \{u_k(x_0)\})$. Set $j = 1$.

Step 1 Given state x_k at time k, find a control sequence $\{u_k(x_j)\}_{k=0}^{N-1} \in \mathcal{U}_N(x_j)$ such that

$$V_N'(x_{j-1}) - \alpha l(u_0(x_{j-1}), x_{j-1}) \geq V_N(x_j, \{u_k(x_j)\}).$$

Set $V_N'(x_j) = V_N(x_j, \{u_k(x_j)\})$.

Step 2 $j \leftarrow j + 1$. Go to step 1.

Lemma 5.5.1 *Suppose* $F(\cdot)$, $\kappa_f(\cdot)$, *and* $\mathcal{X}_f$ *satisfy conditions* **C0-C2**, *assumptions* **A0-A2** *hold, and* $N \geq 1$. *If* $x_0 \in \mathcal{X}_N$, *then, for all* $k \geq 0$, *there exists a control sequence* $\{u_k(x_j)\}_{j=0}^{N-1} \in \mathcal{U}_N(x_j)$ *such that*

$$V_N'(x_{j-1}) - \alpha l(u_0(x_{j-1}), x_{j-1}) \geq V_N(x_j, \{u_k(x_j)\}).$$

where $x_{j+1} = Ax_j + Bu_0(x_j)$.

Proof. The proof is by induction. $x_0 \in \mathcal{X}_N$ implies there exists $\{u_k(x_0)\}_{k=0}^{N-1} \in \mathcal{U}_N(x_0)$. Let $V_N'(x_0) = V_N(x_0, \{u_k(x_0)\})$. Conditions **C0** and **C1** imply

$$\{u_1(x_0), u_2(x_0), \dots, u_{N-1}(x_0), \kappa_f(x_N(x_0))\} \in \mathcal{U}_N(G(x_0)).$$

If we choose

$$\{u_k(x_1)\}_{k=0}^{N-1} = \{u_1(x_0), u_2(x_0), \dots, u_{N-1}(x_0), \kappa_f(x(N; x_0, \{u_j(x_0)\}))\}$$

then condition **C2** implies

$$V_N'(x_0) - V_N'(x_1, \{u_k(x_1)\}_{k=0}^{N-1}) \geq l(u_0(x_0), x_0).$$

Now suppose $\{u_k(x_j)\}_{k=0}^{N-1} \in \mathcal{U}_N(x_j)$ and

$$V_N'(x_{j-1}) - \alpha l(u_0(x_{j-1}), x_{j-1}) \geq V_N(x_j, \{u_k(x_j)\}).$$

If we choose

$$\{u_k(x_{j+1}\}_{k=0}^{N-1} = \{u_1(x_j), u_2(x_j), \dots, u_{N-1}(x_j), \kappa_f(x(N; x_j, \{u_j(x_j)\}))\}$$

then

$$V_N'(x_j) - V_N(x_{j+1}, \{u_k(x_j)\}) \geq l(u_0(x_j), x_j).$$

and the lemma follows as claimed. □

Proposition 5.5.2 *Suppose $F(\cdot)$, $\kappa_f(\cdot)$, and $\mathcal{X}_f$ satisfy conditions* **C0**-**C2**, *assumptions* **A0**-**A2** *hold, $N \geq 1$, the system (5.1) is constrained controllable, and there exists a number $a > 0$, a K-function $\gamma(\cdot)$, and control sequence $\{u_k(z)\}_{k=0}^{N-1}$ such that, for all $z \in N_a$,*

$$\sigma(\|z\|) \geq V_N(z, \{u_k(z)\}).$$

Then suboptimal MPC is a stable control law for the system (5.1) and $G^{(k)}(z) \to 0$ as $k \to \infty$ for all $z \in \mathcal{X}_N$.

Proof. We first establish convergence. By assumption

$$V_N'(x_j) - \alpha l(u_0(x_{j-1}), x_{j-1}) \geq V_N'(x_j) \geq 0.$$

This inequality implies the sequence $\{V_N(x_j)\}$ is nonincreasing. As $V_N(\cdot)$ is bounded below (by assumption **A2**), the sequence $\{V_N(x_j)\}$ converges and

$$G^{(k)}(x_0) \to 0.$$

To prove stability, choose $\delta < \sigma^{-1}(\epsilon)$ and $\tilde{T} = 0$. The existence of $\sigma^{-1}(\cdot)$ follows from Fact 5.3.2. □

Remark 5.5.3 *The condition*

$$\sigma(\|z\|) \geq V_N(z, \{u_k(z)\}).$$

states that the suboptimal value function $V_N(\cdot)$ is continuous at the origin. In other words, if the state x is close to the origin, then the control is small.

5.6 Conclusion

In this chapter we reviewed the basic theory of nonlinear MPC. We limited our discussion to state feedback. As we demonstrated, we can generalize many MPC results using the conditions **C0-C2**. In the subsequent chapters, we discuss constrained linear MPC and associated issues related to target tracking, computation, robustness, and output feedback. While these issues are relevant also in nonlinear MPC, their development is easier for linear systems and many technical issues still need to be resolved in order to apply these results to nonlinear systems.

Issues not discussed included inverse optimality (Magni and Sepulchre 1997) and hybrid control using MPC with mixed-integer nonlinear programming (Slupphaug and Foss 1997, Bemporad and Morari 1999)). The latter is a promising area of research. In most processes, the regulatory control system and the supervisory logic are designed separately, and the interactions between the two levels are handled in an ad-hoc manner. In the hybrid control framework, these two element are combined, thereby allowing one to optimally configure the two regulatory levels (c.f. (Branicky, Borkar and Mitter 1998)) The hybrid control framework allows also for discrete controls, such as on/off values and logical overrides, and prioritized objectives and constraints. However, these problems are combinatorial, and the online solution of these optimization problems may not always be feasible.

Chapter 6

Steady States and Constraints in Linear Model Predictive Control[1]

6.1 Introduction

Model predictive control (MPC) is an optimization based strategy that uses a plant model to predict the effect of potential control action on the evolving state of the plant. At each time step, an open-loop optimal control problem is solved and the input profile is injected into the plant until a new measurement becomes available. The updated plant information is used to formulate and solve a new open-loop optimal control problem.

Since MPC is formulated as an optimization problem, inequality constraints are a natural addition to the controller. The ability to handle explicitly input and output constraints may be viewed as one of the major factors for the success of MPC in process control. Operation at constraints is so common that it may be regarded as the rule rather than the exception in chemical process operations. Consider the classic example of temperature control of an exothermic reactor. In order to maximize profit, one may wish to maximize reactor feed rate. At some feed rate, however, the cooling capacity reaches a constraint. As disturbances, such as heat exchanger fouling occur, the feed rate is manipulated to maximize production with some safety margin while maintaining cooling capacity at its constraint. If a disturbance were to decrease the reactor feed temperature, however, then the cooling rate would be decreased so the reaction would not extinguish. So in many practical situations of this type, inputs, cooling rate in this example, are maintained at constraints in the normal steady-state operation. The main purpose of this chapter is to extend the existing linear MPC theory to handle this important industrial case.

While constraints improve the appeal of MPC as an advanced control strategy, they complicate the implementation of the controller. In addition to the computational burden, constraints necessitate additional safeguards to guarantee that the controller is stabilizing. One method to guarantee nominal stability is to formulate the model predictive controller on an infinite horizon (Keerthi and Gilbert 1988). Infinite horizon formulations are appealing because, for the nominal case, the predicted open-loop and the achieved closed-loop responses are identical and the effect of tuning parameters is, therefore, more intuitive.

In this chapter we formulate MPC as an infinite horizon optimal control strategy with a quadratic performance criterion. We use the following discrete time model of the plant

$$x_{j+1} = Ax_j + B(u_j + d), \tag{6.1a}$$

$$y_j = Cx_j + p, \tag{6.1b}$$

[1] **This chapter was published in essentially the same form as Rao and Rawlings (1999)**

where $x_j \in \mathbb{R}^n$ is the state vector, $u_j \in \mathbb{R}^m$ is the input vector, and $y_j \in \mathbb{R}^q$ is the output vector. $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{q\times m}$ are, respectively, the state transition matrix, the input distribution matrix, and the measurement matrix. The subscript $j \in \mathbb{I}_+$ denotes the discrete time sampling instant. The affine terms $d \in \mathbb{R}^m$ and $p \in \mathbb{R}^q$ serve the purpose of adding integral control. They may be interpreted as modeling the effect of constant disturbances influencing the input and output, respectively. Muske and Rawlings (1993) provide a discussion of how to estimate p and d. Assuming that the state of the plant is perfectly measured, we define model predictive control as the feedback law $u_j = g(x_j)$ that minimizes

$$\Phi = \frac{1}{2}\sum_{j=0}^{\infty}(y_j - \bar{y})^T Q(y_j - \bar{y}) + (u_j - \bar{u})^T R(u_j - \bar{u}) + \Delta u_j^T S \Delta u_j, \tag{6.2}$$

where $\Delta u_j \triangleq u_j - u_{j-1}$. The matrices Q, R, and S are assumed to be symmetric positive definite. The vector $\bar{y}$ is the desired output target and $\bar{u}$ is the desired input target, assumed for simplicity to be time invariant. When the complete state of the plant is not measured, as is almost always the case, the addition of a state estimator is necessary. Since state estimation is beyond the scope of this article, we assume that the control and estimation problems can be separated.

The steady-state aspect of the control problem is to determine appropriate values of $(y_{\text{ss}}, x_{\text{ss}}, u_{\text{ss}})$ satisfying the following relation

$$x_{\text{ss}} = Ax_{\text{ss}} + B(u_{\text{ss}} + d), \tag{6.3a}$$

$$y_{\text{ss}} = Cx_{\text{ss}} + p. \tag{6.3b}$$

Ideally, $y_{\text{ss}} = \bar{y}$ and $u_{\text{ss}} = \bar{u}$. However, process limitations and constraints may prevent the system from reaching the desired steady state. The goal of the target calculation is to find the feasible triple $(y_{\text{ss}}, x_{\text{ss}}, u_{\text{ss}})$ such that y_{ss} and u_{ss} are as close as possible to $\bar{y}$ and $\bar{u}$. We address the target calculation in Section 6.2.

To simplify the analysis and formulation, we transform (6.2) using deviation variables to the generic infinite horizon quadratic criterion

$$\Phi = \frac{1}{2}\sum_{j=0}^{\infty} z_j^T Q z_j + v_j^T R v_j + \Delta v_j^T S \Delta v_j. \tag{6.4}$$

The original criterion (6.2) can be recovered from (6.4) by making the following substitutions:

$$z_j \leftarrow y_j - Cx_{\text{ss}} - p, \quad w_j \leftarrow x_j - x_{\text{ss}}, \quad v_j \leftarrow u_j - u_{\text{ss}}.$$

By using deviation variables we treat separately the steady-state and the dynamic elements of the control problem, thereby simplifying the overall analysis of the controller.

The dynamic aspect of the control problem is to control (y, x, u) to the steady-state values $(y_{\text{ss}}, x_{\text{ss}}, u_{\text{ss}})$ in the face of constraints, which may be active at the steady-state operating point. This part of the problem is discussed in Section 6.3. In particular we determine the state feedback law $v_j = \rho(w_j)$ that minimizes (6.4). When there are no inequality constraints, the feedback law is the linear quadratic regulator. However, with the addition of inequality constraints, there may not exist an analytic form for $\rho(w_j)$. In such cases where an analytic solution is unavailable, the feedback law is obtained by repetitively solving the open-loop optimal control problem. This strategy allows us to consider only the encountered sequence of measured states rather than the entire state space. For a further discussion, see Mayne (1995*a*).

If we consider only linear constraints on the input, input velocity, and outputs of the form

$$u_{\min} \leq Du_k \leq u_{\max}, \qquad -\Delta_u \leq \Delta u_k \leq \Delta_u, \qquad y_{\min} \leq Cx_k \leq y_{\max}, \tag{6.5}$$

where $D \in \mathbb{R}^{n_D \times m}$ and $C \in \mathbb{R}^{n_C \times q}$, we formulate the regulator as the solution to the following infinite horizon optimal control problem

$$\min_{\{w_k, v_k\}} \Phi(x_j) = \frac{1}{2} \sum_{k=0}^{\infty} z_k^T Q z_k + v_k^T R v_k + \Delta v_k^T S \Delta v_k, \tag{6.6}$$

subject to the constraints

$$w_0 = x_j - x_{\text{ss}}, \qquad v_{-1} = u_{j-1} - u_{\text{ss}}, \qquad w_{k+1} = A w_k + B v_k, \qquad z_k = C w_k \tag{6.7a}$$

$$u_{\min} - u_{\text{ss}} \leq D v_k \leq u_{\max} - u_{\text{ss}}, \quad -\Delta_u \leq \Delta v_k \leq \Delta_u, \tag{6.7b}$$

$$y_{\min} - y_{\text{ss}} \leq C w_k \leq y_{\max} - y_{\text{ss}}. \tag{6.7c}$$

If we denote

$$\{w_{k+1}^*(x_j), v_k^*(x_j)\}_{k=0}^{\infty} = \arg\min \Phi(x_j),$$

then the control law is

$$\rho(x_j) = v_0^*(x_j).$$

We address the regulation problem in Section 6.3.

Combining the solution of the target tracking problem and the constrained regulator, we define the MPC algorithm as follows:

1. Obtain an estimate of the state and disturbances $\Longrightarrow (x_j, p, d)$
2. Determine the steady-state target $\Longrightarrow (y_{\text{ss}}, x_{\text{ss}}, u_{\text{ss}})$
3. Solve the regulation problem $\Longrightarrow v_j$
4. Let $u_j = v_j + u_{\text{ss}}$
5. Set $j \leftarrow j + 1$. Go to step 1.

6.2 Target Calculation

When the number of the inputs equals the number of outputs, the solution to the unconstrained target problem is obtained using the steady-state gain matrix, assuming such a matrix exists (i.e. the system has no integrators). However for systems with unequal numbers of inputs and outputs, integrators, or inequality constraints, the target calculation is formulated as a mathematical program (Muske and Rawlings 1993, Muske 1997). When there are at least as many inputs as outputs, multiple combinations of inputs may yield the desired output target at steady state. For such systems, a mathematical program with a least squares objective is formulated to determine the best combinations of inputs. When the number of outputs is greater than the number of inputs, situations exist in which no combination of inputs satisfies the output target at steady state. For such cases, we formulate a mathematical program that determines the steady-state output $y_{\text{ss}} \neq \bar{y}$ that is closest to $\bar{y}$ in a least squares sense.

Instead of solving separate problems to establish the target, we prefer to solve one problem for both situations. Through the use of an exact penalty (Fletcher 1987), we formulate the target tracking problem as a single quadratic program that achieves the output target if possible, and relaxes the problem in a l_1/l_2^2 optimal sense if the target is infeasible. We formulate the soft constraint

$$\begin{aligned} \bar{y} - C x_{\text{ss}} - p &\leq \eta, \\ \bar{y} - C x_{\text{ss}} - p &\geq -\eta, \\ \eta &\geq 0, \end{aligned} \tag{6.8}$$

by relaxing the constraint $Cx_{ss} = \bar{y}$ using the slack variable η. By suitably penalizing η, we guarantee that the relaxed constraint is binding when it is feasible. We formulate the exact soft constraint by adding an l_1/l_2^2 penalty to the objective function. The l_1/l_2^2 penalty is simply the combination of a linear penalty $q_{ss}^T\eta$ and a quadratic penalty $\eta^T Q_{ss}\eta$, where the elements of q_{ss} are strictly non-negative and Q_{ss} is a symmetric positive definite matrix. By choosing the linear penalty sufficiently large, the soft constraint is guaranteed to be exact. A lower bound on the elements of q_{ss} to ensure that the original hard constraints are satisfied by the solution cannot be calculated explicitly without knowing the solution to the original problem, because the lower bound depends on the optimal Lagrange multipliers for the original problem. In theory, a conservative state-dependent upper bound for these multipliers may be obtained by exploiting the Lipschitz continuity of the quadratic program (Hager 1979). However, in practice, we rarely need to guarantee that the l_1/l_2^2 penalty is exact. Rather, we use approximate values for q_{ss} obtained by computational experience. In terms of constructing an exact penalty, the quadratic term is superfluous. However, the quadratic term adds an extra degree of freedom for tuning and is necessary to guarantee uniqueness.

We now formulate the target tracking optimization as the following quadratic program

$$\min_{x_{ss}, u_{ss}, \eta} \frac{1}{2}\left(\eta^T Q_{ss}\eta + (u_{ss} - \bar{u})^T R_{ss}(u_{ss} - \bar{u})\right) + q_{ss}^T\eta \tag{6.9}$$

subject to the constraints

$$\begin{bmatrix} I-A & -B & 0 \\ C & 0 & I \\ C & 0 & -I \end{bmatrix} \begin{bmatrix} x_{ss} \\ u_{ss} \\ \eta \end{bmatrix} \left\{ \begin{matrix} = \\ \geq \\ \leq \end{matrix} \right\} \begin{bmatrix} Bd \\ \bar{y}-p \\ \bar{y}-p \end{bmatrix}, \qquad \eta \geq 0, \tag{6.10a}$$

$$u_{\min} \leq Du_{ss} \leq u_{\max}, \qquad y_{\min} \leq Cx_{ss} + p \leq y_{\max}, \tag{6.10b}$$

where R_{ss} and Q_{ss} are assumed to be symmetric positive definite.

Because x_{ss} is not explicitly in the objective function, the question arises as to whether the solution to Equation 6.9 is unique. If the feasible region is non-empty, the solution exists because the quadratic program is bounded below on the feasible region (Frank and Wolfe 1956). If Q_{ss} and R_{ss} are symmetric positive definite, η and u_{ss} are uniquely determined by the solution of the quadratic program. However, without a quadratic penalty on x_{ss}, there is no guarantee that the resulting solution for x_{ss} is unique. Non–uniqueness in the steady-state value of x_{ss} presents potential problems for the controller, because the origin of the regulator is not fixed at each sample time. Consider, for example, a tank where the level is unmeasured (i.e. an unobservable integrator). The steady-state solution is to set $u_{ss} = 0$ (i.e. balance the flows). However, any level x_{ss}, within bounds, is an optimal alternative. Likewise, at the next time instant, a different level also would be a suitably optimal steady-state target. The resulting closed-loop performance for the system could be erratic, because the controller may constantly adjust the level of the tank, never letting the system settle to a steady state.

In order to avoid such situations, we restrict our discussion to detectable systems, and recommend redesign if a system does not meet this assumption. For detectable systems, the solution to the quadratic program is unique assuming the feasible region is nonempty. The details of the proof are given in Appendix 6.5.1. Uniqueness is also guaranteed when only the integrators are observable. For the practitioner this condition translates into the requirement that all levels are measured. The reason we choose the stronger condition of detectability is that if good control is desired, then the unstable modes of the system should be observable. Detectability is also required to guarantee nominal stability of the regulator.

Empty feasible regions are a result of the inequality constraints (6.10b). Without the inequality constraints (6.10b) the feasible region is nonempty, thereby guaranteeing the existence of a feasible and unique solution under the condition of detectability. For example, the solution $(u_{ss}, x_{ss}, \eta) = (-d, 0, |\bar{y} -$

$p|$) is feasible. However, the addition of the inequality constraints (6.10b) presents the possibility of infeasibility. Even with well-defined constraints, $u_{\text{min}} < u_{\text{max}}$ and $y_{\text{min}} < y_{\text{max}}$, disturbances may render the feasible region empty. Since the constraints on the input usually result from physical limitations such as valve saturation, relaxing only the output constraints is one possibility to circumvent infeasibilities. Assuming that $u_{\text{min}} \leq -d \leq u_{\text{max}}$, the feasible region is always nonempty. However, we contend that the output constraints should not be relaxed in the target calculation. Rather, an infeasible solution, readily determined during the initial phase in the solution of the quadratic program, should be used as an indicator of a process exception. While relaxing the output constraints in the dynamic regulator is common practice (Ricker et al. 1988, Genceli and Nikolaou 1993, de Oliveira and Biegler 1994, Zheng and Morari 1995, Scokaert and Rawlings 1999), the output constraint violations are transient. By relaxing output constraints in the target calculation on the other hand, the controller seeks a steady-state target that continuously violates the output constraints. The steady violation indicates that the controller is unable to compensate adequately for the disturbance and, therefore, should indicate a process exception.

6.3 Receding Horizon Regulator

Because our implementation of dynamic control in the presence of active steady-state constraints employs an infinite horizon, we first briefly discuss the solution to infinite horizon problems.

6.3.1 Infinite Horizon Optimal Control Problem

Given the calculated steady state we formulate the regulator as the following infinite horizon optimal control problem

$$\min_{\{w_k, v_k\}} \Phi(x_j) = \frac{1}{2} \sum_{k=0}^{\infty} w_k^T C^T Q C w_k + v_k^T R v_k + \Delta v_k S \Delta v_k, \tag{6.11}$$

subject to the constraints

$$w_0 = x_j - x_{\text{ss}}, \qquad v_{-1} = u_{j-1} - u_{\text{ss}}, \qquad w_{k+1} = A w_k + B v_k, \tag{6.12a}$$

$$u_{\text{min}} - u_{\text{ss}} \leq D v_k \leq u_{\text{max}} - u_{\text{ss}}, \qquad -\Delta_u \leq \Delta v_k \leq \Delta_u, \tag{6.12b}$$

$$\bar{y} - y_{\text{ss}} \leq C w_k \leq \bar{y} - y_{\text{ss}}. \tag{6.12c}$$

We assume that Q and R are symmetric positive definite matrices. We also assume that the origin,$(w_j, v_j) = (0,0)$, is an element of the feasible region $\mathbb{W} \times \mathbb{V}$ [2]. If the pair (A, B) is stabilizable, the pair $(A, Q^{1/2}C)$ is detectable, and a solution exists to (6.11)–(6.12), then $x_j = 0$ is an exponentially stable fixed point of the closed-loop system (Scokaert and Rawlings 1996).

For unstable state transition matrices, the direct solution of (6.11)–(6.12) is ill-conditioned, because the system dynamics are propagated through the unstable A matrix. To improve the conditioning of the optimization, we reparameterize the input as $v_k = Lw_k + r_k$, where L is a linear stabilizing feedback gain for (A, B) (Keerthi 1986, Rossiter, Rice and Kouvaritakis 1997). The system model becomes

$$w_{k+1} = (A + BL) w_k + B r_k, \tag{6.13}$$

where r_k is the new input. By initially specifying a stabilizing, potentially infeasible, trajectory, we can improve the numerical conditioning of the optimization by propagating the system dynamics through the stable $(A + BL)$ matrix.

---

[2] $\mathbb{W} = \{w |\ y_{\text{min}} - y_{\text{ss}} \leq Cw \leq y_{\text{max}} - y_{\text{ss}}\}$, $\mathbb{V} = \{v |\ u_{\text{min}} \leq Dv \leq u_{\text{max}},\ -\Delta_u - u_{\text{ss}} \leq \Delta v \leq \Delta_u - u_{\text{ss}}\}$

By expanding Δv_k and substituting in for v_k, we transform (6.11)–(6.12) into the following more tractable form:

$$\min_{\{w_k, v_k\}} \Phi(x_j) = \frac{1}{2}\sum_{k=0}^{\infty}(w_k^T Q w_k + v_k^T R v_k + 2 w_k^T M v_k), \tag{6.14}$$

subject to the following constraints:

$$w_0 = x_j, \qquad w_{k+1} = A w_k + B v_k, \tag{6.15a}$$

$$d_{\text{min}} \le D v_k - G w_k \le d_{\text{max}}, \qquad y_{\text{min}} - y_{\text{ss}} \le C w_k \le y_{\text{max}} - y_{\text{ss}}. \tag{6.15b}$$

The original formulation (6.11)–(6.12) can be recovered from (6.14)–(6.15) by making the following substitutions into the second formulation:

$$x_j \leftarrow \begin{bmatrix} x_j - x_{\text{ss}} \\ u_{j-1} - u_{\text{ss}} \end{bmatrix}, \; w_k \leftarrow \begin{bmatrix} w_k \\ v_{k-1} \end{bmatrix}, \; v_k \leftarrow r_k, \; A \leftarrow \begin{bmatrix} A + BL & 0 \\ L & 0 \end{bmatrix},$$

$$B \leftarrow \begin{bmatrix} B \\ I \end{bmatrix}, \; Q \leftarrow \begin{bmatrix} C^T Q C + L^T (R+S) L & -L^T S \\ -SL & S \end{bmatrix}, \; M \leftarrow \begin{bmatrix} L^T (R+S) \\ -S \end{bmatrix},$$

$$R \leftarrow R + S, \; D \leftarrow \begin{bmatrix} D \\ I \end{bmatrix}, \; G \leftarrow \begin{bmatrix} -DL & 0 \\ -L & I \end{bmatrix},$$

$$d_{\text{max}} \leftarrow \begin{bmatrix} u_{\text{max}} - u_{\text{ss}} \\ \Delta_u \end{bmatrix}, \; d_{\text{min}} \leftarrow \begin{bmatrix} u_{\text{min}} - u_{\text{ss}} \\ -\Delta_u \end{bmatrix}, \; C \leftarrow \begin{bmatrix} C & 0 \end{bmatrix}.$$

While the formulation (6.14)–(6.15) is theoretically appealing, the solution is intractable in its current form, because it is necessary to consider an infinite number of decision variables. In order to obtain a computationally tractable formulation, we reformulate the optimization in a finite dimensional decision space.

Several authors have considered this problem in various forms. In this chapter, we concentrate on the constrained linear quadratic methods proposed in the literature (Keerthi 1986, Sznaier and Damborg 1987, Chmielewski and Manousiouthakis 1996, Scokaert and Rawlings 1996, Scokaert and Rawlings 1998). The key concept behind these methods is to recognize that the inequality constraints remain active only for a finite number of sample steps along the prediction horizon. We demonstrate informally this concept as follows: if we assume that there exists a feasible solution to (6.14), (6.15), then the state and input trajectories $\{w_k, v_k\}_{k=0}^{\infty}$ approach the origin exponentially. Furthermore, if we assume the origin is contained within the interior of the feasible region $\mathbb{W} \times \mathbb{V}$ (we address the case where the origin lies on the boundary of the feasible region in the next section), then there exists a positively invariant convex set (Gilbert and Tan 1991)

$$\mathcal{O}_\infty = \{w | \; (A + BK)^j w \in \mathbb{W}_K, \; \forall j \ge 0\} \tag{6.16}$$

such that the optimal unconstrained feedback law $v = Kw$ is feasible for all future time. The set $\mathbb{W}_K$ is the feasible region projected onto the state space by the linear control K (i.e. $\mathbb{W}_K = \{w|(w, Kw) \in \mathbb{W} \times \mathbb{V}\}$). Because the state and input trajectories approach the origin exponentially, there exists a finite $N^\star$ such that the state trajectory $\{w_k\}_{k=N^\star}^{\infty}$ is contained in $\mathcal{O}_\infty$.

In order to guarantee that the inequality constraints (6.15b) are satisfied on the infinite horizon, $N^\star$ must be chosen such that $w_{N^\star} \in \mathcal{O}_\infty$. Since the value of $N^\star$ depends on x_j, we need to account for the variable decision horizon length in the optimization. We formulate the variable horizon length regulator as the following optimization

$$\min_{\{w_k, v_k, N\}} \Phi(x_j) = \frac{1}{2}\sum_{k=0}^{N-1}(w_k^T Q w_k + v_k^T R v_k + 2 w_k^T M v_k) + \frac{1}{2} w_N^T \Pi w_N, \tag{6.17}$$

subject to the constraints

$$w_0 = x_j, \qquad w_{k+1} = Aw_k + Bv_k, \tag{6.18a}$$

$$d_{\min} \le Dv_k - Gw_k \le d_{\max}, \qquad y_{\min} - y_{\mathrm{ss}} \le Cw_k \le y_{\max} - y_{\mathrm{ss}}, \tag{6.18b}$$

$$w_N \in \mathcal{O}_\infty. \tag{6.18c}$$

The cost to go Π is determined from the discrete-time algebraic Riccati equation

$$\Pi = A^T \Pi A + Q - (A^T \Pi B + M)(R + B^T \Pi B)^{-1}(B^T \Pi A + M^T), \tag{6.19}$$

for which many reliable solution algorithms exist. The variable horizon formulation is similar to the dual-mode receding horizon controller (Michalska and Mayne 1993) for nonlinear system with the linear quadratic regulator chosen as the stabilizing linear controller.

While the problem (6.17)–(6.18) is formulated on a finite horizon, the solution cannot, in general, be obtained in real-time since the problem is a mixed-integer program. Rather than try to solve directly (6.17)–(6.18), we address the problem of determining N^* from a variety of semi-implicit schemes while maintaining the quadratic programming structure in the subsequent optimizations.

Gilbert and Tan (1991) show that there exist a finite number t^* such that $\mathcal{O}_{t^*}$ is equivalent to the maximal $\mathcal{O}_\infty$, where

$$\mathcal{O}_t = \{w| \ \ (A + BK)^j w \in \mathbb{W}_K\,, \ \ \text{for } j = 0, \ldots, t\ \}. \tag{6.20}$$

They also present an algorithm for determining t^* that is formulated efficiently as a finite number of linear programs. Their method provides an easy check whether, for a fixed N, the solution to (6.17)–(6.18) is feasible (i.e. $w_N \in \mathcal{O}_\infty$). The check consists of determining whether state and input trajectories generated by unconstrained control law $v_k = Kw_k$ from the initial condition w_N are feasible with respect to inequality constraints for t^* time steps in the future. If the check fails, then the optimization (6.17)–(6.18) needs to be resolved with a longer control horizon $N' > N$ since $w_N \notin \mathcal{O}_\infty$. The process is repeated until $w_{N'} \in \mathcal{O}_\infty$.

When the set of initial conditions $\{w_0\}$ is compact, Chmielewski and Manousiouthakis (1996) present a method for calculating an upper bound $\bar{N}$ on N^* using bounding arguments on the optimal cost function Φ^*. Given a set $\mathbb{P} = \{x^1, \ldots, x^m\}$ of initial conditions, the optimal cost function $\Phi^*(x)$ is a convex function defined on the convex hull (co) of $\mathbb{P}$. An upper bound $\bar{\Phi}(x)$ on the optimal cost $\Phi^*(x)$ for $x \in \mathrm{co}(\mathbb{P})$ is obtained by the corresponding convex combinations of optimal cost functions $\Phi^*(x^j)$ for $x^j \in \mathbb{P}$. The upper bound on N^* is obtained by recognizing that the state trajectory w_j only remains outside of $\mathcal{O}_\infty$ for a finite number of stages. A lower bound q on the cost of $w_j^T Q w_j$ can be generated for $x_j \notin \mathcal{O}_\infty$ (see (Chmielewski and Manousiouthakis 1996) for explicit details). It then follows that $N^* \le \bar{\Phi}(x)/q$. Further refinement of the upper bound can be obtained by including the terminal stage penalty Π in the analysis.

When a bound on the initial conditions w_0 is known *a priori*, calculating an upper bound $\bar{N}$ is appealing, because one need not iteratively determine N^* online. However, generating this bound *a priori* requires significant process knowledge. Changing operating conditions and disturbances may lead to initial conditions that violate any previously specified bound. In such cases, we again need to determine N^* online. Furthermore, the decision of how to construct the basis for $\mathbb{P}$ is complicated, since the number of points increases exponentially in higher dimensions. Even when a bound is available and a logical basis is constructed, the upper bounds are often conservative, as demonstrated in the following example.

Example 6.3.1 *Comparison of online and offline determination of* N^*

Consider the regulation of the following double integrator system

$$\dot{x}_1 = x_2, \tag{6.21a}$$
$$\dot{x}_2 = u, \tag{6.21b}$$

sampled at a frequency of 10 Hertz with $y = x_1$ and the input constraint $|u| \leq 1$. For $Q = 1$, $R = 1$, $S = 0$, and the initial condition $x_0 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$. For this initial condition, $N^* = 13$ was required to guarantee that the constraints are satisfied on the infinite horizon.

The Chmielewski and Manousiouthakis method generates a least upper bound of 361 for N^*. This value was determined using the true infinite horizon cost for $x_0 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$. In practice, only an upper bound on the cost is available for the infinite horizon cost, so the upper bound on N^* is often greater than the least upper bound for N^*. We can compare these results with the repetitive strategy where N is increased until $w_N \in \mathcal{O}_\infty$. Since there exist algorithms whose computational cost is $O(N)$ (Rao, Wright and Rawlings 1998), we can expect that the computational cost is approximately a linear function of N. If $N = 1$ initially and the control horizon is increased by unit steps, then the total computational cost is approximately $91 \times C$, where C is the computational cost required solve the optimization for $N = 1$. If we increase the horizon geometrically with a factor of 2 as advocated by Scokaert and Rawlings (1998), then the total computational cost is approximately $31 \times C$. In practice, larger initial values of N are used. A good heuristic is to choose initially $N = t^*$. For this example, $t^* = 15$. As the example demonstrates, the online determination is significantly less computationally expensive than the offline determination. Furthermore, for the online determination, we can bound the computational cost by $4N^* \times C$, for this example $52 \times C$, when we increase the horizon geometrically with a factor of 2 (Scokaert and Rawlings 1998). With the offline determination, we have no bounds on the computational cost (other than it is finite), and, as the example demonstrates, a computational effort an order of magnitude greater than required is possible. Therefore, we suggest the use of the iterative, online determination for N^*. $\diamond$

6.3.2 Feasibility and Soft Constraints

In the formulation of the MPC problem, some state constraints are imposed by physical limitations such as valve saturation. Other constraints are less important; they may represent desired ranges of operation for the plant, for instance. In some situations, no set of inputs and states for the MPC problem may satisfy all of these constraints. Rather than having the algorithm declare infeasibility and return without a result, we prefer a solution that enforces some constraints strictly ("hard constraints"), while relaxing others and replacing them with penalties on their violation ("soft constraints"). This problem has been discussed by numerous authors (Ricker et al. 1988, Genceli and Nikolaou 1993, de Oliveira and Biegler 1994, Zheng and Morari 1995, Scokaert and Rawlings 1999),

Scokaert and Rawlings (1999) replace the soft constraints with penalty terms in the objective that are a combination of ℓ_1 norms and squared ℓ_2 norms of the constraint violations. Assuming for simplicity that all state constraints $y_{\min} - y_{ss} \leq Cx_k \leq y_{\max} - y_{ss}$ in (6.18b) are softened in this way, we obtain the following modification to the objective (6.17):

$$\min_{u,x,\epsilon} \Phi(u, x, \epsilon) = \frac{1}{2} \sum_{k=0}^{N-1} (x_k^T Q x_k + u_k^T R u_k + 2x_k^T M u_k + \epsilon_k^T Z \epsilon_k) + z^T \epsilon_k + \frac{1}{2} x_N^T \bar{P} x_N, \tag{6.22}$$

where the constraint violations ϵ_k are defined by the following formulae (which replace $y_{\min} - y_{\mathrm{ss}} \leq Cx_k \leq y_{\max} - y_{\mathrm{ss}}$):

$$Cx_k - \epsilon_k \leq y_{\max} - y_{\mathrm{ss}}, \tag{6.23a}$$

$$Cx_k + \epsilon_k \geq y_{\min} - y_{\mathrm{ss}}, \tag{6.23b}$$

$$\epsilon_k \geq 0 \tag{6.23c}$$

and the elements of the vector z are nonnegative, while Z is an symmetric positive semi-definite matrix. It is known that when the weighting z on the ℓ_1 terms is sufficiently large (see, for example, Section 12.3 in Fletcher (1987)), and when the original problem (6.17)-(6.18) has a nonempty feasible region, then local minimizers of problem (6.17)-(6.18) modified by (6.22)-(6.23) defined above correspond to local solutions of the unmodified problem (6.17)- (6.18). Under these conditions, (6.22) together with the constraints (6.23) is referred to as an exact penalty formulation of the original objective (6.17) with the original constraints $y_{\min} - y_{\mathrm{ss}} \leq Cx_k \leq y_{\max} - y_{\mathrm{ss}}$. This formulation has the advantage that it can still yield a solution when the original problem (6.17)-(6.18) is infeasible.

Prior to actually solving the problem, we cannot know how large the elements of z must be chosen to make the exact penalty property hold. (The threshold value depends on the optimal multipliers for the original problem (6.17)-(6.18).) A conservative state-dependent upper bound for these multipliers can be obtained by exploiting the Lipschitz continuity of the quadratic program (Hager 1979). In practice, however, the exact penalty is not critical, since by definition soft constraints need not be satisfied exactly. Reasonable controller performance can often be achieved by setting $z = 0$ and choosing Z to be a positive diagonal matrix. In fact, the inclusion of the ℓ_2 term $\epsilon_k^T Z \epsilon_k$ is not needed at all for the exact penalty property to hold, but is included here to provide a little more flexibility in the modeling.

6.3.3 Boundary Solutions and Suboptimal Approximations

All papers on constrained linear MPC include the assumption that the origin lies in the interior of the feasible region (Keerthi and Gilbert 1988, Sznaier and Damborg 1987, Rawlings and Muske 1993, Chmielewski and Manousiouthakis 1996, Scokaert and Rawlings 1996). However, as Section 6.2 indicates, this assumption is often violated. In practice, one often encounters situations in which a valve saturates or a control variable rides at a performance constraint during steady-state operation. In these situations, the origin is on the boundary of the feasible region. Table 6.1 lists all of the examples that are discussed in the chapter and summarizes the main points illustrated with each example. Consider the following example.

Example 6.3.2 *Saturating inputs at steady state*

Prett and Morari (1987) presented the following model

$$G(s) = \begin{bmatrix} \frac{4.05e^{-27s}}{50s+1} & \frac{1.77e^{-28s}}{60s+1} & \frac{5.88e^{-27s}}{50s+1} \\ \frac{5.39e^{-18s}}{50s+1} & \frac{5.72e^{-14s}}{60s+1} & \frac{6.90e^{-15s}}{40s+1} \\ \frac{4.38e^{-20s}}{33s+1} & \frac{4.42e^{-22s}}{44s+1} & \frac{7.20}{19s+1} \end{bmatrix} \tag{6.24}$$

for a heavy oil fractionator as the benchmark process for the Shell standard control problem. The three inputs of the process represent the product draw rate from the top of the column (u_1), the product draw rate from the side of the column (u_2), and the reflux heat duty for the bottom of the column (u_3). The three outputs of the process represent the draw composition (y_1) from the top of the column, the draw

Example	Description
1	A comparison of methods for calculating N^* in solving infinite horizon problems
2	Steady-state inputs on boundary for Shell problem
3	Steady-state outputs on boundary for Furnace problem
4	Example of system whose input never settles on constraint or remains in interior of feasible region
5	Simple numerical example of input constrained regulator
5	Dynamic response of Shell problem subject to setpoint change
6	Dynamic response of Shell problem subject to disturbance
7	Example of output constrained regulator: endpoint constraint necessary
8	Example of output constrained regulator: boundary solution

Table 6.1: Brief Synopsis of Examples

composition (y_2) from the side of the column, and the reflux temperature at the bottom of the column (y_3). Prett and Garcia also present the following disturbance model

$$G_d(s) = \begin{bmatrix} \frac{1.20e^{-27s}}{45s+1} & \frac{1.44e^{-27s}}{40s+1} \\ \frac{1.52e^{-15s}}{25s+1} & \frac{1.83e^{-15s}}{20s+1} \\ \frac{1.14}{27s+1} & \frac{1.26}{32s+1} \end{bmatrix} \tag{6.25}$$

for the heavy oil fractionator. The two disturbances are the reflux heat duty for the intermediate section of the column (d_1) and the reflux heat duty for the top of the column (d_2). Both models were sampled with a period of 4 minutes.

The inputs are constrained between -0.5 and 0.5. An input velocity constraint of 0.20 is also imposed. In addition to constraints on the inputs, the outputs are constrained between -0.5 and 0.5. The following tuning parameters were chosen: $Q(Q_{ss}) = I$ and $R(R_{ss}) = I$.

Since the origin is shifted by the steady-state target calculation, output target changes and measured disturbances may force the origin to lie on the boundary of the feasible region. An example of an output target change that causes the inputs to saturate at steady state is

$$\bar{y} = \begin{bmatrix} 0.3 \\ 0.3 \\ -0.3 \end{bmatrix} \Longrightarrow u_{ss} = \begin{bmatrix} 0.5 \\ -0.1 \\ -0.26 \end{bmatrix}, \quad y_{ss} = \begin{bmatrix} 0.3 \\ 0.3 \\ -0.15 \end{bmatrix}. \tag{6.26}$$

Note that since the input (u_1) saturates, the system is unable to attain the desired target (i.e. $\bar{y} \neq y_{ss}$). Figure 6.1 illustrates how the input constraints constrain the attainable region of the output space. Likewise, an example of a step disturbance, d_{step}, that causes the inputs to saturate at steady state is

$$d_{step} = \begin{bmatrix} 0.6 \\ 0.5 \end{bmatrix} \Longrightarrow u_{ss} = \begin{bmatrix} -0.5 \\ 0.04 \\ 0.09 \end{bmatrix}. \tag{6.27}$$

Steady-state outputs at performance constraints are a consequence of choosing an output target at ± 0.5 or choosing an infeasible target. ◇

Example 6.3.3 *Constrained outputs at steady state*

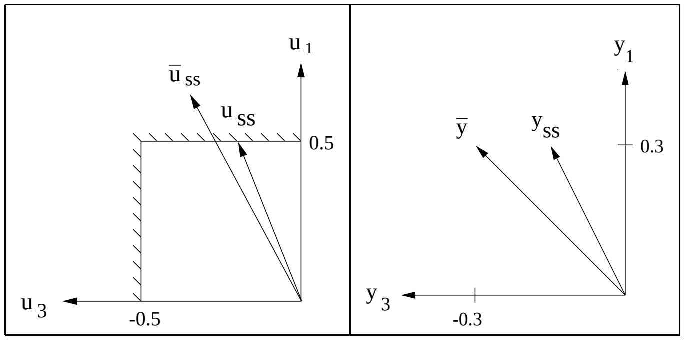


Figure 6.1: The limiting effect of the input constraints on the ability to attain the desired output target for Example 6.3.2. The vector $\bar{u}_{\text{ss}}$ denotes the unconstrained input required to achieve the target $\bar{y}$.

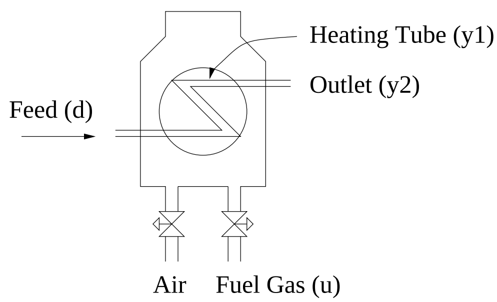


Figure 6.2: Preheater furnace

Consider the control of a furnace depicted in Figure 6.2, where the objective is to preheat the feed to a desired output temperature. The input variable is the fuel gas flowrate. In addition to the input constraint caused by valve saturation, there is a maximum limit for the furnace temperature in order to prevent the furnace tubes from melting. While temporary violations of the constraint are tolerable, long term violations are not. If we assume that the heating tube temperature and the heat transfer are linear functions of the fuel gas flowrate, then a simplified, steady-state energy balance neglecting heat loss yields the following dimensionless model for the furnace system

$$y_1 = \alpha u, \tag{6.28a}$$

$$y_2 = \beta u + d, \tag{6.28b}$$

in which y_1 is the heating tube temperature, y_2 is the outlet temperature, d is the inlet temperature, and u is the fuel gas flowrate. For simplicity we scale the variable such that $\alpha = \beta = 1$. Assume nominal conditions are $u = 8$, $d = 5$, and $\bar{y} = \begin{bmatrix} 8 & 13 \end{bmatrix}^T$ and the maximum limit for the furnace temperature is $y_1 \leq 10$. Suppose that there is an upstream disturbance that causes the inlet temperature to drop to $d = 2$. To compensate for the disturbance, the fuel gas flow rate would have to increase to 11 for the output temperature (y_2) to remain at its target. However, the furnace temperature constraint would allow the flowrate to increase only to 10. The resulting steady-state output obtained with $Q_{ss} = 1$, $R_{ss} = 1$, $q_{ss} = 100$, and $\bar{u} = 0$ is $y_{\text{ss}} = \begin{bmatrix} 10 & 12 \end{bmatrix}^T$, which lies on the boundary of the feasible region. Figure 6.3 depicts the effect of the disturbance on the attainable region of the output space. $\diamond$

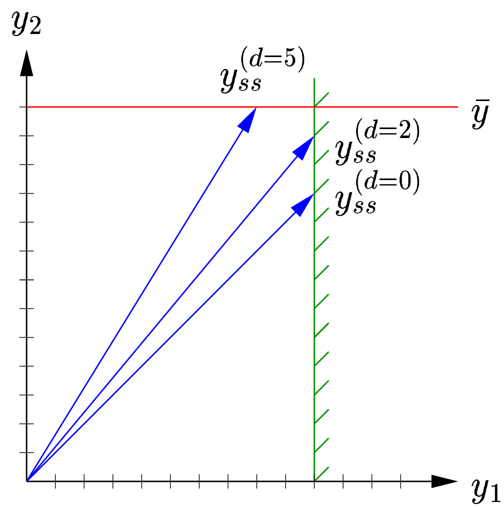


Figure 6.3: The inability of the outlet temperature to reach the desired steady-state target due to the tube temperature constraint in Example 6.3.3.

These situations complicate the formulation of the infinite horizon optimization, because $\mathcal{O}_\infty$ does not exist for all such systems controlled with the unconstrained optimal feedback regulator. Figure 6.4 displays some of the potential input and state trajectory characteristics that are possible when the origin lies on the boundary of the feasible region.

An example of a system that displays the first characteristic is a stable first order system with initial conditions in the interior of the feasible region. An example of a system displaying the second characteristic is given in the following example.

Example 6.3.4 *A system where the control does not become permanently active or inactive on the constraint*

Consider the following system

$$w_{k+1} = \begin{bmatrix} 0.5477 & 0.8208 & 0 \\ -0.8208 & 0.5067 & 0 \\ 0 & 0 & 0.8 \end{bmatrix} w_k, + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} v_k, \tag{6.29a}$$

$$y_k = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} w_k, \tag{6.29b}$$

subject to the input constraint $v_k \leq 0$. Figure 6.5 details the optimal input profile subject to the initial disturbance $w_0 = \begin{bmatrix} 3 & 3 & 0 \end{bmatrix}^T$ with tuning parameters $Q = 1$ and $R = 1$. While the input trajectory converges towards the origin (see Figure 6.5), numerical calculations indicate that the input does not become active on the constraint or stay strictly in the interior of the feasible region. $\diamond$

Examples of systems displaying the third characteristic are given in Examples 6.3.7 and 6.3.8. While the first trajectory offers the possibility of constructing $\mathcal{O}_\infty$, for the second and third trajectory we are unable to construct $\mathcal{O}_\infty$ with a finite number of inequality constraints, because the constraints remains active for infinite time. Each of the possibilities could be handled individually. However, the task of segregating their behavior *a priori* is difficult. To circumvent this problem, we approximate (6.17)–(6.18) by restricting the evolution of the input and state trajectories generated by the linear control law to the null space of the active constraints at the origin. This suboptimal strategy coincides with

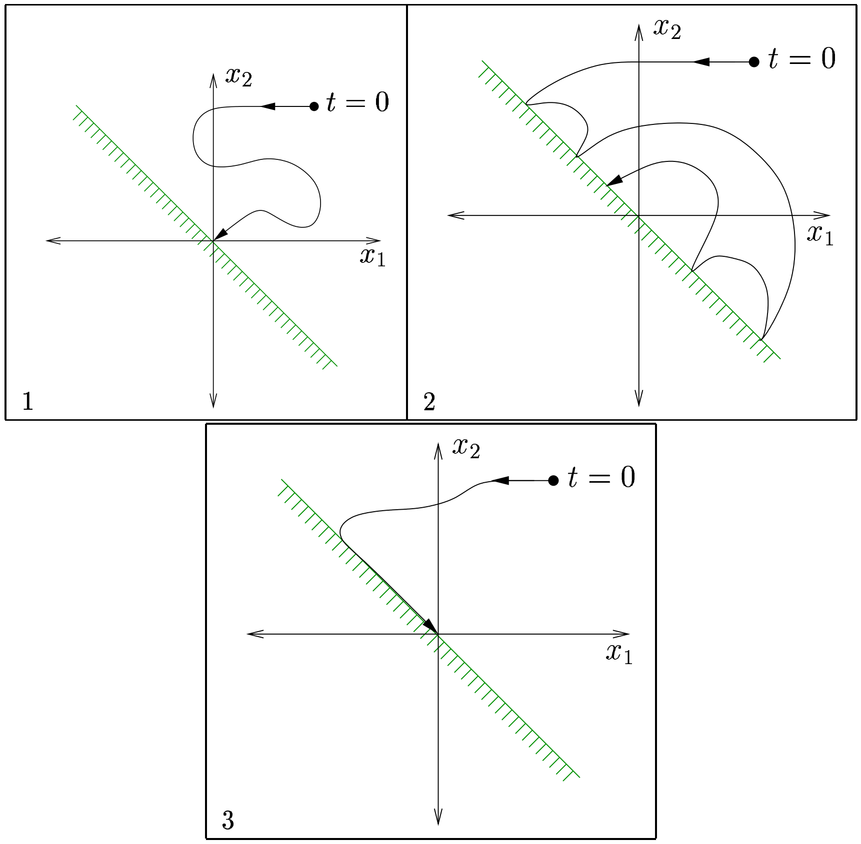


Figure 6.4: Some potential input and state trajectories when the origin is on the boundary of the feasible region. To prevent problems associated with trajectories 1 and 2, the proposed algorithm forces the closed-loop response to adhere to a trajectory similar to 3.

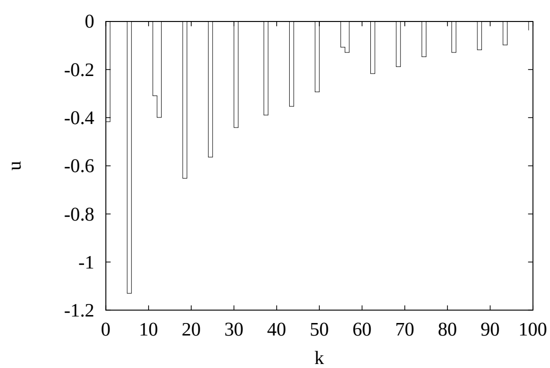


Figure 6.5: Input Trajectory for Example 6.3.4

forcing the state and input trajectories to adhere to the third path depicted in Figure 6.4. By forcing the trajectory of system onto the constraints active at the origin, we guarantee the existence of an $\mathcal{O}_\infty$, because the invariant set needs to account only for the constraints inactive at the origin. The constraints active at the origin are feasible by construction.

We accomplish the boundary approximation by constructing the optimal linear feedback law $\bar{K}$ that constrains evolution of the closed-loop system to constraints active at the origin. The set $\mathcal{O}_\infty$ is constructed as before with the following differences: the new linear control law is $\bar{K}$ and the inequality constraints active at the origin are discarded. The infinite horizon regulator is constructed as the solution of (6.17)–(6.18) with the new $\mathcal{O}_\infty$ and the cost to go $\bar{\Pi}$ associated with $\bar{K}$. We treat the situation of input and state constraints separately.

Active Input Constraints We recast the problem for handling input constraints active at the origin as finding the linear feedback controller that minimizes the infinite horizon quadratic objective (6.14) subject to the equality constraints [3]

$$w_{k+1} = Aw_k + Bv_k, \tag{6.30a}$$
$$\bar{D}v_k = 0, \tag{6.30b}$$

where the over-bar denotes the subset of the inequality constraints, $d_{\min} \leq Dv_k \leq d_{\max}$, that are active at the origin. The matrix $\bar{D} \in \mathbb{R}^{n_{\bar{D}} \times m}$, where $n_{\bar{D}}$ is equal to the number of inputs with constraints active at the origin. The state dependence of the input constraints due to (6.13) can be reformulated solely in terms of the input v_k by removing the parameterization $v_k = Lw_k + r_k$ for $k \geq N$. We derive the linear optimal controller as follows. We first define the operator

$$\mathcal{K} : (A, B, Q, R, M) \to (K, \Pi), \tag{6.31}$$

where K is the linear gain for the optimal unconstrained regulator and Π is the solution to the associated Riccati equation. If we let $\mathcal{N}_{\bar{D}}$ be an orthonormal basis for the null space of $\bar{D}$, then $p_k^{\bar{D}} = \mathcal{N}_{\bar{D}}^T v_k$ represents the input projected to the null space $\mathcal{N}_{\bar{D}}$. Because the equality constraint is feasible for all $p_k^{\bar{D}}$, we can substitute for v_k in the state equation and the objective function yielding the following expression

$$(K_{\bar{D}}, \bar{\Pi}) = \mathcal{K}(A, B\mathcal{N}_{\bar{D}}, Q, \mathcal{N}_{\bar{D}}^T R \mathcal{N}_{\bar{D}}, M\mathcal{N}_{\bar{D}}) \tag{6.32}$$

for the solution of the constrained feedback law. If $(A, B\mathcal{N}_{\bar{D}})$ is stabilizable, then

$$\bar{K} \triangleq \mathcal{N}_{\bar{D}} K_{\bar{D}}. \tag{6.33}$$

If $(A, B\mathcal{N}_{\bar{D}})$ is not stabilizable, we need to zero the modes of the system that are both uncontrollable and unstable at $k = N^\star$ in order to guarantee nominal stability. By first performing a Kalman decomposition to construct a basis for the uncontrollable subspace, a basis for the corresponding uncontrollable and unstable modes is constructed using either a Jordan or Schur decomposition. We remark that if $\bar{D}$ is full rank, then it is necessary to zero the inputs at the end of the control horizon. The regulator then reduces to the one discussed by Rawlings and Muske (1993). In the following two examples we show the closed-loop response of the heavy oil fractionator with the output target change and disturbance described in Example 6.3.2.

Example 6.3.5 *Heavy oil fractionator: closed-loop response subject to output target change*

Consider the closed-loop response of the heavy oil fractionator described in Example 6.3.2 subject to the output target change described in (6.26). Figure 6.6 shows the closed-loop response subject to the

[3] We do not need to account for $\bar{G}$ because the velocity constraints are not active at steady state.

output target change. As discussed in Example 6.3.2, the input constraints prevent the system from attaining the desired target $\bar{y}$ at steady state. Instead the controller seeks a target that causes the top draw (u_1) to saturate at its upper limit. Figure 6.6 shows that, after initially saturating, the top draw asymptotically approaches its upper limit as the closed-loop system settles at its specified steady-state values. While the open-loop trajectory of the controller specifies that the top draw saturates at $k = N^\star$, the receding horizon aspect of the regulator allows for an asymptotic approach. This feedback effect diminishes the performance degradation due to the boundary projection in the actual closed-loop. $\diamond$

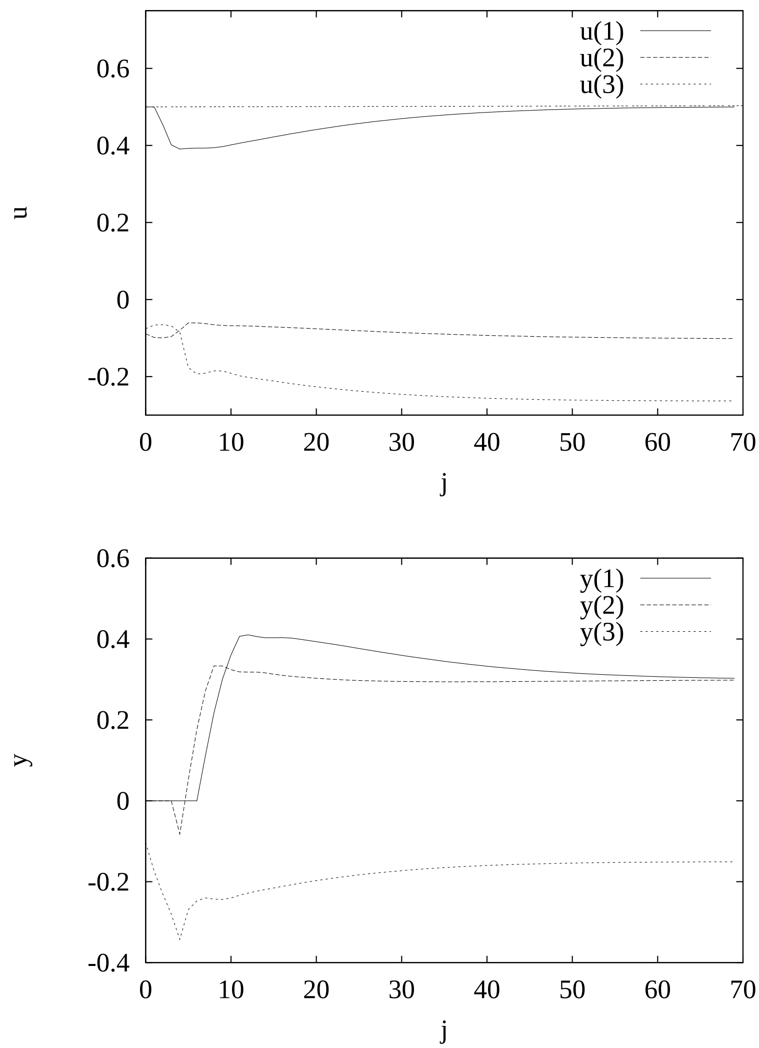


Figure 6.6: Closed-loop Response for Example 6.3.5

Example 6.3.6 *Heavy oil fractionator: closed-loop response subject to a disturbance*

Consider the closed-loop response of the heavy oil fractionator described in Example 6.3.2 subject to the disturbance described in (6.27). An output disturbance model was used to detect the disturbance (Muske and Rawlings 1993). Figure 6.7 displays the closed-loop response subject to the disturbance. To reject the disturbance, the controller seeks an input target that causes the top draw (u_1) to saturate at its lower limit. Once again, Figure 6.7 shows that the top draw asymptotically approaches its lower limit as the closed-loop system settles at its specified steady state. In addition, the disturbance causes the output constraints to become infeasible. In order to handle the output infeasibilities in the regulator, the constraints were relaxed using a l_1/l_2^2 exact penalty in the manner described by Scokaert and Rawlings (1996). The output constraints were relaxed with an l_1 penalty of $z = 1000 * e$ and a l_2^2 penalty of $Z = I$, where $e = [1 \ldots 1]^T$. $\diamond$

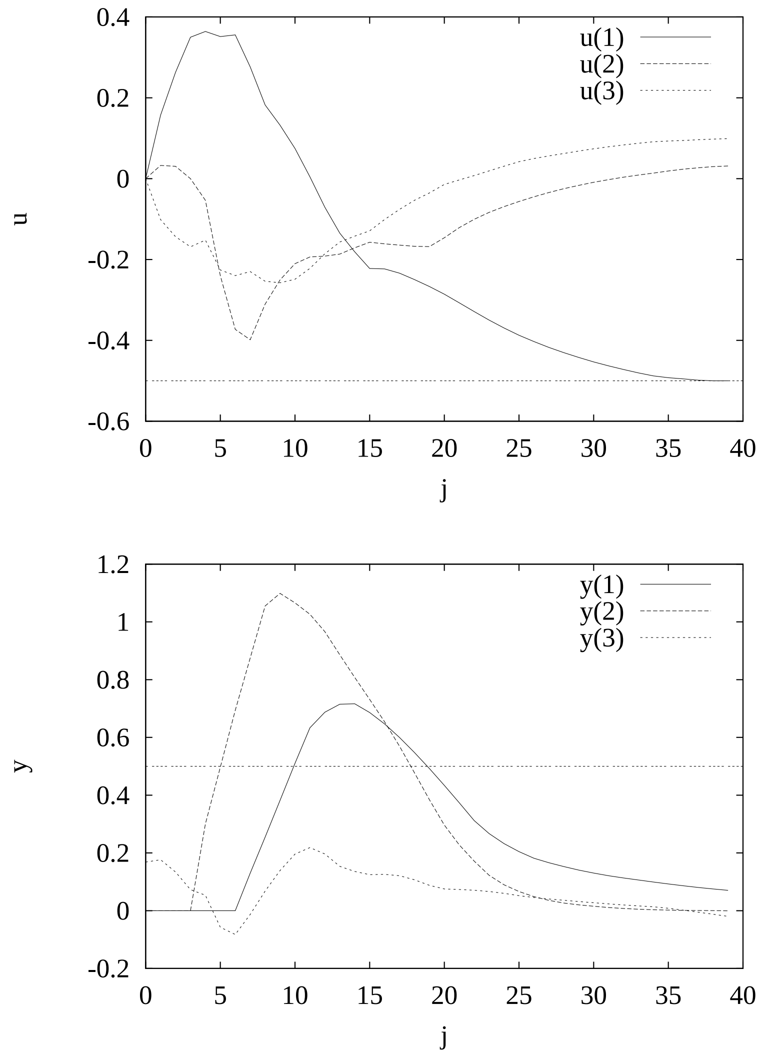


Figure 6.7: Closed-loop Response for Example 6.3.6

Active State Constraints In an analogous manner, the problem of handling state constraints can be reformulated as finding the linear feedback controller that minimizes the infinite horizon quadratic objective (6.14) subject to the constraints

$$w_{k+1} = Aw_k + Bv_k, \tag{6.34a}$$
$$\bar{C}w_k = 0. \tag{6.34b}$$

Unlike the previous situation, there does not always exist a linear feedback regulator that satisfies the state constraints for all k. For such a regulator to exist, we require that the null space of $\bar{C}$ is (A, B) invariant. The definition of (A, B) invariance along with the sufficient conditions for the existence of a regulator is given in Appendix 6.5.2. The condition of (A, B) invariance essentially requires that there are enough degrees of freedom in the input to constrain the evolution of the system to a particular subspace. A system whose uncontrollable modes are observable in the null space of $\bar{C}$ is not (A, B) invariant if the associated basis for the uncontrollable modes is not contained completely in the null space of $\bar{C}$. Assuming that the null space of $\bar{C}$ is (A, B) invariant, then there exists a linear feedback law

$$v_k = L_{\bar{C}}w_k + \mathcal{N}_{B\bar{C}}p_k^{\bar{C}} \tag{6.35}$$

that constrains the system to the null-space of $\bar{C}$ for all $p_k^{\bar{C}}$. The details of the construction are given in Appendix 6.5.2. We obtain the optimal feedback law from the following expression

$$\begin{aligned}(K_{\bar{C}}, \bar{\Pi}) \quad = \quad & \mathcal{K}((A + BL_{\bar{C}}), B\mathcal{N}_{\bar{C}B}, Q + L_{\bar{C}}^T RL_{\bar{C}} + 2ML_{\bar{C}}, \\ & \mathcal{N}_{\bar{C}B}^T R\mathcal{N}_{\bar{C}B}, M\mathcal{N}_{\bar{C}B} + L_{\bar{C}}^T R\mathcal{N}_{\bar{C}B})\end{aligned} \tag{6.36}$$

by substituting in for v_k with the feedback law (6.35). If $((A + BL_{\bar{C}}, B\mathcal{N}_{CB})$ is stabilizable, then

$$\bar{K} \triangleq \mathcal{N}_{\bar{C}B}K_{\bar{C}}. \tag{6.37}$$

Otherwise it is necessary to zero the unstable and uncontrollable modes at $k = N^\star$ in an analogous manner to the input constrained regulator. The boundary approximation to (6.17)–(6.18) is obtained by adding the constraint

$$\bar{C}w_N = 0, \tag{6.38}$$

and calculating $\bar{\Pi}$ using (6.36). In the following two examples, we illustrate that a system must possess excess degrees of freedom in the input for a stabilizing boundary approximation.

Example 6.3.7 *Output constrained regulator with no excess degrees of freedom*

Consider the regulation of the following non-minimum phase system

$$y(s) = \frac{s - 3}{3s^2 + 4s + 2}u(s) \tag{6.39}$$

sampled at a frequency of 10 Hertz subject to the constraint $y_k \geq 0$ and the tuning parameters $Q = 1$ and $R = 1$. One state space realization for this system in discrete time is:

$$A = \begin{bmatrix} 0.9968 & 0.0935 \\ -0.0623 & 0.8721 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0048 \\ 0.0935 \end{bmatrix}, \tag{6.40a}$$
$$C = \begin{bmatrix} -1.0000 & 0.3333 \end{bmatrix}. \tag{6.40b}$$

The system is (A, B) invariant with respect to the null space of C. However, the closed-loop system with the invariant feedback law

$$K_{\bar{C}} = \begin{bmatrix} 38.5608 & -7.4723 \end{bmatrix} \tag{6.41}$$

is unstable. Since there are no additional degrees of freedom, it is necessary to enforce the endpoint constraint $w_N = 0$ for the boundary approximation. Figure 6.8 shows a comparison of the closed-loop response for the constrained regulator with $N = 10$ and the unconstrained regulator with an initial state disturbance of $[-1, 2]^T$. While the output response for the constrained regulator displayed in Figure 6.8 appears better than the unconstrained output response, notice that the input action is far more aggressive for the constrained regulator. $\diamond$

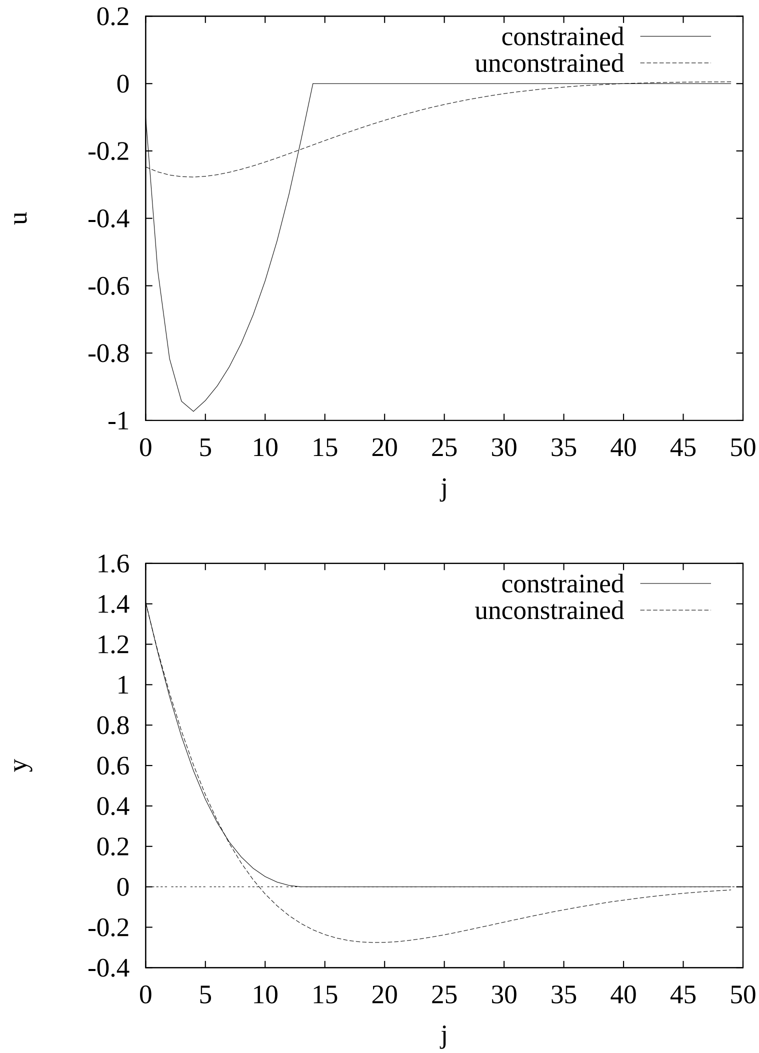


Figure 6.8: Comparison of Closed-loop Responses for Example 6.3.7

Example 6.3.8 *Output constrained regulator with excess degrees of freedom*

Reconsider the regulation of the system in Example 6.3.7 with an additional input

$$y(s) = \frac{s-3}{3s^2+4s+2}u_1(s) + \frac{2}{3s^2+4s+2}u_2(s) \tag{6.42}$$

sampled at a frequency of 10 Hertz subject to the constraints $|u_k| \leq 6$ and $y_k \geq 0$ and the tuning parameters $Q = 1$ and $R = I$.

One state space realization for this system in discrete time is:

$$A = \begin{bmatrix} 0.8877 & -0.0346 \\ 0.1194 & 0.9812 \end{bmatrix}, \quad B = \begin{bmatrix} -0.0827 & 0.0395 \\ 0.0212 & 0.0026 \end{bmatrix},$$
$$C = \begin{bmatrix} 0 & 1.2472 \end{bmatrix}.$$

The system is again (A, B) invariant with respect to the null space of C. In contrast, with the addition of the extra degree of freedom, the closed-loop system with the invariant feedback law

$$K_{\bar{C}} = \begin{bmatrix} -5.5486 & -45.5972 \\ -0.6805 & -5.5921 \end{bmatrix} \tag{6.43}$$

is stabilizable. Figure 6.9 shows a comparison of the closed-loop responses for the constrained regulator with $N = 9$ and the unconstrained regulator subject to the initial condition is $x_0 = [1, \ -3]^T$. Figure 6.10 shows the phase portraits of the closed-loop responses for the constrained and unconstrained regulator. Once again the output response for the constrained regulator appears superior to the unconstrained output response. Notice, however, that the input is far more aggressive for the constrained regulator.$\diamond$

The combined problem of both state and input constraints is solved by reconsidering (6.36) after making the following substitutions:

$$B \leftarrow B\mathcal{N}_{\bar{D}}, \ \ R \leftarrow \mathcal{N}_{\bar{D}}^T R \mathcal{N}_{\bar{D}}, \ \ M \leftarrow M\mathcal{N}_{\bar{D}}. \tag{6.44}$$

It is not difficult to prove the proposed control algorithm is asymptotically stable. Convergence of the regulator is straightforward to demonstrate using standard arguments (for example, see (Keerthi and Gilbert 1988)). Establishing nominal stability is more subtle than the usual arguments. In particular, the definition of stability needs to be adjusted to account for perturbations only in the feasible region.

6.4 Concluding Remarks

The main contribution of this chapter has been to establish techniques for handling inequality constraints active at steady state, a case that has not been treated in previous model predictive control (MPC) theory. Through a series of examples, we show how this case is significant in applications.

As an alternative to the approach outlined in this chapter, one could consider moving any inequality constraint line that passes through the origin a small distance away from the origin, after which existing theory would apply. Choosing this distance is problematic, however. If a small distance is chosen, the output admissible set may be small, and the required horizon may be large and the on-line computation is inefficient. If a somewhat larger distance is chosen, the economic performance of the plant suffers because the steady-state target is no longer close to the true plant constraints. Of course, one could always avoid the issue by using a finite horizon and terminal constraint, but that choice is not as good as the approach outlined here.

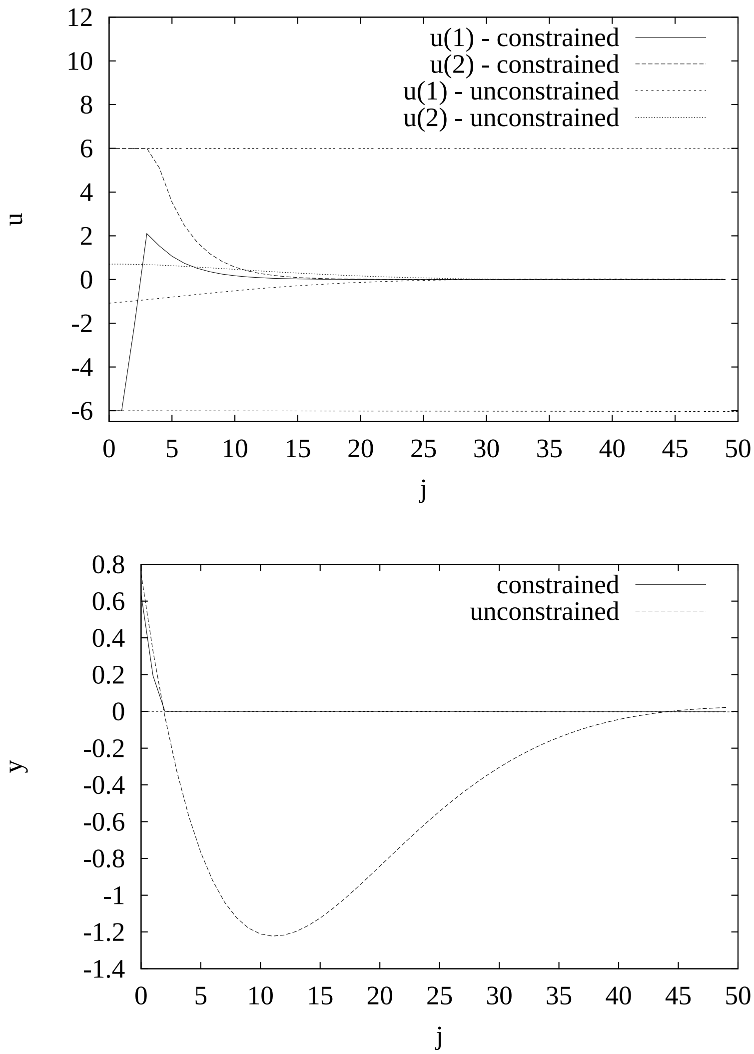


Figure 6.9: Comparison of Closed-loop Responses for Example 6.3.8

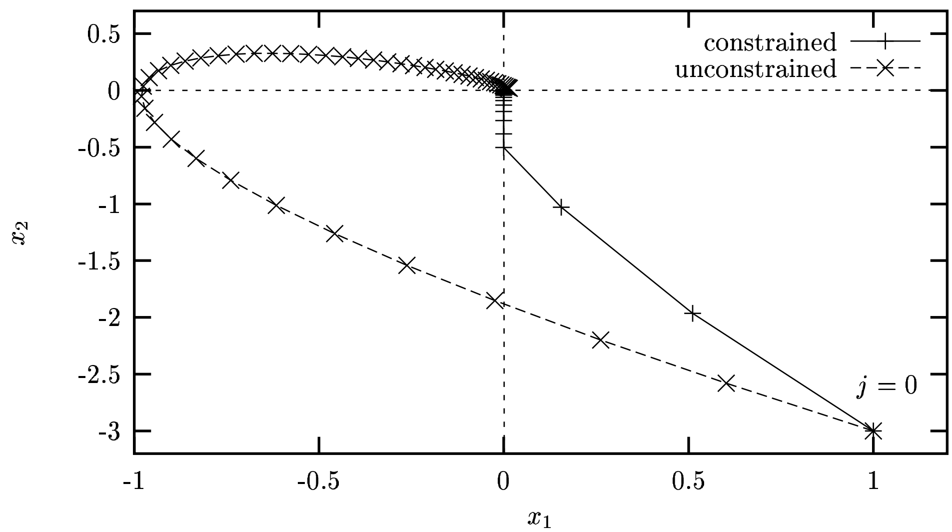


Figure 6.10: Closed-loop Phase Portraits for Example 6.3.8. The initial condition is $x_0 = [1,\ -3]^T$.

6.5 Appendix

6.5.1 Proof of Uniqueness for Target Calculation

Without loss of generality, we ignore the inequality constraints on the decision variables and consider only the constraints given by (6.10a).

Theorem 6.5.1 *If $Q_{\text{ss}}, R_{\text{ss}} > 0$, $q_{ss} > 0$, and (A, C) is detectable, then the solution to (6.9) subject to the constraints given by (6.10a) is unique.*

Proof. Using the Hautus lemma (Sontag 1990), detectability implies the following rank condition

$$\text{rank}\left(\mathcal{H} \triangleq \begin{bmatrix} \lambda I - A \\ C \\ C \end{bmatrix}\right) = n \tag{6.45}$$

for all $\lambda \in \mathbb{C}$ with magnitude greater than or equal to 1. It is sufficient to consider only the extended Hautus matrix, $\mathcal{H}$, with $\lambda = 1$. Since rank $\mathcal{H} = n$, $\mathcal{H}$ has full column rank. It then follows that x_{ss} is uniquely determined from the following equation

$$\mathcal{H}x_{\text{ss}} = \begin{bmatrix} B(u_{\text{ss}} + d) \\ \bar{y} - p - (\eta - t_1) \\ \bar{y} - p + (\eta - t_2) \end{bmatrix} \tag{6.46}$$

where t_1 and t_2 are positive slacks for the inequality constraints (6.10a). If $\bar{x}_{\text{ss}} \neq x_{\text{ss}}$ is another solution, then there necessarily exists another solution $(\bar{u}_{\text{ss}}, \bar{\eta} - \bar{t}_{\{1,2\}}) \neq (u_{\text{ss}}, \eta - t_{\{1,2\}})$. Since the positive slack accounts for $\eta = |\bar{y} - Cx_{\text{ss}} - p|$, η is uniquely determined by $(\eta - t_{\{1,2\}})$. However, since the objective function is a strictly convex function of u_{ss} and η, $\bar{x}_{\text{ss}}$ cannot be another solution without contradicting optimality. □

Remark 6.5.2 *An additional consequence of a unique target is that the target calculation is stable to perturbations. Since the quadratic program is continuous in a point-to-set topology, uniqueness of the target guarantees that the solution is continuous in a point-to-point topology (Berge 1963).*

6.5.2 State Constrained Linear Quadratic Regulator

In this section, we describe sufficient conditions for the construction of a state constrained linear feedback controller. The key concept is (A, B) invariance with respect to the null space of $\bar{C}$. For further details of (A, B) invariance with respect to an arbitrary subspace, see Section 4.3 of (Sontag 1990).

Definition 6.5.3 *The null space of $\bar{C}$ is (A, B) invariant if and only if $\forall w$, $\exists v$ such that $\bar{C}w = 0$ implies $\bar{C}(Aw + Bv) = 0$.*

Theorem 6.5.4 *The null space of $\bar{C}$ is (A, B) invariant if and only if there exists a linear feedback control law that constrains the evolution of (A, B) to the subspace $\bar{C}w = 0$.*

The sufficiency is immediate (Sontag 1990). Before proving necessity of Theorem 6.5.4, we first derive necessary and sufficient conditions for the null space of $\bar{C}$ to be (A, B) invariant. Let $\mathcal{N}_{\bar{C}}$ be an orthonormal basis for the null-space of $\bar{C}$ and let $\zeta = -\mathcal{N}_{\bar{C}}^T w$. We recast the constraints (6.34) as

$$\bar{C}A\mathcal{N}_{\bar{C}}\zeta = \bar{C}Bv. \tag{6.47}$$

Let $\mathcal{R}_{(\cdot)}$ denote the range space of $(\cdot)$.

Lemma 6.5.5 *The null space of $\bar{C}$ is (A,B) invariant if and only if $\mathcal{R}_{\bar{C}A\mathcal{N}_{\bar{C}}} \subseteq \mathcal{R}_{\bar{C}B}$.*

Proof. Suppose first that $\mathcal{R}_{\bar{C}A\mathcal{N}_{\bar{C}}} \subseteq \mathcal{R}_{\bar{C}B}$. It follows directly that $\forall \zeta$, $\exists v$ such that $\bar{C}A\mathcal{N}_{\bar{C}}\zeta = \bar{C}Bv$ since the column space of $B\bar{C}$ contains the columns space of $\bar{C}A\mathcal{N}_{\bar{C}}$. Hence, $\bar{C}$ is (A,B) invariant as claimed.

Now suppose that the null space of $\bar{C}$ is (A,B) invariant. Then $\forall \zeta$, $\exists v$ such that $\bar{C}A\mathcal{N}_{\bar{C}}\zeta = \bar{C}Bv$. Therefore, by definition, $\mathcal{R}_{\bar{C}A\mathcal{N}_{\bar{C}}} \subseteq \mathcal{R}_{\bar{C}B}$ as claimed. $\square$

Let O_A and O_B denote the orthonormal bases for the columns spaces of $\bar{C}A\mathcal{N}_{\bar{C}}$ and $\bar{C}B$ respectively.

Corollary 6.5.6 *The null space of $\bar{C}$ is (A,B) invariant if and only if $(O_B O_B^T - I)O_A = 0$.*

Proof. An equivalent condition for $\mathcal{R}_{\bar{C}A\mathcal{N}_{\bar{C}}} \subseteq \mathcal{R}_{\bar{C}B}$ is that $\forall \zeta$, $\exists v$ such that $O_A\zeta = O_B v$. Solving for v yields $v = O_B^T O_A \zeta$. Direct substitution yields the desired result. $\square$

Proof. [Theorem 6.5.4] We construct a feedback law that constrains w to the null-space of $\bar{C}$ by decomposing the operator $\bar{C}B$ into its range space and null spaces yielding $v = -(\bar{C}B)^+ CAw + \mathcal{N}_{\bar{C}B} p^{\bar{C}}$ where $p^{\bar{C}}$ is the input projected to the null space of $\bar{C}B$ and $(\cdot)^+$ denotes the pseudo-inverse. If $K_{\bar{C}} \triangleq -(\bar{C}B)^+ CA$, then we construct the linear feedback law $v = Kw + \mathcal{N}_{\bar{C}B} p^{\bar{C}}$ as claimed. $\square$

Remark 6.5.7 *The vector $p^{\bar{C}}$ represents the excess degrees of freedom in the inputs with respect to the constraint $\bar{C}w = 0$. Since the control law constrains the system to the subspace $\bar{C}w = 0$ for all $p^{\bar{C}}$, we construct the optimal state constrained regulator by first constructing a linear quadratic regulator, K, for the system $(A + BK_{\bar{C}}, B\mathcal{N}_{\bar{C}B})$. The full state constrained regulator is obtained by combining the two linear regulators as follows*

$$v = (K_{\bar{C}} + \mathcal{N}_{\bar{C}B} K)\, w. \tag{6.48}$$

Chapter 7

Linear Programming and Model Predictive Control [1]

7.1 Introduction

Traditionally model predictive control has been formulated using a quadratic criterion. Part of the popularity of the quadratic criterion from a theoretical standpoint is due to its mathematical convenience. From a numerical standpoint, the quadratic criterion is popular, because the resulting optimization can be cast as a quadratic program. For the unconstrained case, the linear quadratic optimal control problem is solved efficiently using dynamic programming. This solution technique has the desirable property that the computational cost scales linearly in the horizon length N as opposed to cubically for the general least squares solution. While the addition of constraints negates the possibility of a general analytic solution to the optimal control problem, the quadratic program may be structured in an analogous manner to the unconstrained problem, yielding linear growth in the horizon length N. Approaches to structuring the optimal control problem with a linear quadratic objective utilizing sparse matrix methods are available in the literature (Wright 1993, Biegler 1997, Rao et al. 1998). See Chapter 8 for a discussion of sparse matrix methods.

Recently Dave, Willig, Kudva, Pekny and Doyle (1997) have advocated the use of an l_1/l_∞ norm as a performance criterion for MPC. One motivation is that it allows us to formulate the optimal control problem as a linear problem. Obtaining solutions to linear program is less computationally demanding than obtaining a solution to a quadratic program of the same size and complexity, so it may be preferable to formulate MPC as a linear program. The concept of using linear programming is not new and has been considered by many authors in optimal control (e.g. (Zadeh and Whalen 1962, Outraka 1976)) and in MPC (e.g. (Propoi 1963, Chang and Seborg 1983, Morshedi, Cutler and Skrovanek 1985, Keerthi and Gilbert 1986, Campo and Morari 1986, Campo and Morari 1987, Campo and Morari 1989, Allwright and Papavasiliou 1992, Genceli and Nikolaou 1993)). A review of some MPC research with non-quadratic objectives can be found in the paper by García, Prett and Morari (1989). The main theoretical objection to linear programming formulations is that analytic solutions are generally unavailable due to the nonsmoothness of the objective function. The nonsmoothness is one of the prime reasons why stability analysis for linear programming formulations has been lacking. Notable exceptions include the works of Keerthi and Gilbert (1988), who use an endpoint constraint, Genceli and Nikolaou (1993), who consider finite impulse response models, and Shamma and Xiong (1997), who provide a numerical test whether a given horizon is sufficiently long to guarantee stability for unconstrained MPC.

In this chapter we examine linear programming formulations of MPC. We begin our discussion by

[1] **Portions of the chapter were published in Rao and Rawlings (1998*b*) and Rao and Rawlings (2000)**

presenting in Section 7.2 a stabilizing formulation of MPC with a general l_p criterion. In Section 7.3 we analyze the qualitative properties of MPC with an l_1 criterion. Unlike MPC with a quadratic criterion, the choice of the tuning parameters for the l_1 formulation may result in appreciably different closed-loop performance. In particular, we demonstrate how the nonsmoothness of the objective may yield either dead-beat or idle control performance.

7.2 Stabilizing MPC with l_p Criterion

Consider the regulation following linear discrete-time representation of the plant

$$x_{k+1} = Ax_k + Bu_k, \quad k \geq 0, \tag{7.1a}$$

$$y_k = Cx_k \tag{7.1b}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, and $y_k \in \mathbb{R}^q$. We formulate the regulator as the feedback law $\eta(\hat{x}_j)$ that generates the sequence $\{u_k\}_{k=0}^{\infty}$, where $\eta(\hat{x}_j) \triangleq u_0$, that minimizes the infinite horizon objective function

$$\Phi(\hat{x}_k) = \sum_{k=0}^{\infty} \|\bar{R}u_k\|_p + \|\bar{Q}y_k\|_p, \tag{7.2}$$

subject to (7.1), the initial condition $x_0 = \hat{x}_j$, and the constraints

$$u_{\min} \leq Du_k \leq u_{\max}, \tag{7.3a}$$

$$-\Delta_u \leq \Delta u_k \leq \Delta_u, \tag{7.3b}$$

$$y_{\min} \leq y_k \leq y_{\max}, \tag{7.3c}$$

where

$$\|x\|_p := \left(\sum_{i=1}^{n} |x^{(i)}|^p \right)^{1/p}$$

and $x^{(i)}$ denotes the i^{th} entry of the vector x. Common examples of l_p norms are the sum norm (l_1 norm)

$$\|x\|_1 \triangleq |x^{(1)}| + \ldots + |x^{(n)}|$$

and the max norm (l_∞ norm)

$$\|x\|_\infty \triangleq \max\{|x^{(1)}|, \ldots, |x^{(n)}|\}.$$

The vector $\hat{x}_j$ denotes the current state estimate of the plant at time index j. By suitably adjusting the origin, the regulator can account for target tracking and disturbance rejection (Muske and Rawlings 1993). We make the following assumptions: a) (A, B) is stabilizable and (C, A) is detectable; b) $\bar{Q}$ and $\bar{R}$ are diagonal matrices with positive elements; and c) the origin $(u_k, x_k) = 0$ is contained within the interior of the feasible region (7.3). If a feasible solution exists, then the origin is an asymptotically stable fixed point for the feedback controller (Keerthi and Gilbert 1988).

With the notable exceptions discussed in Section 7.3, analytic solutions to (7.2) are generally unavailable, because the l_p norm has a kink at the origin (see Figure 7.1). To circumvent the computational barrier imposed by the infinite horizon calculation, we employ a stable finite horizon approximation. Our method is analogous to the technique employed by Rawlings and Muske (1993) for a quadratic criterion.

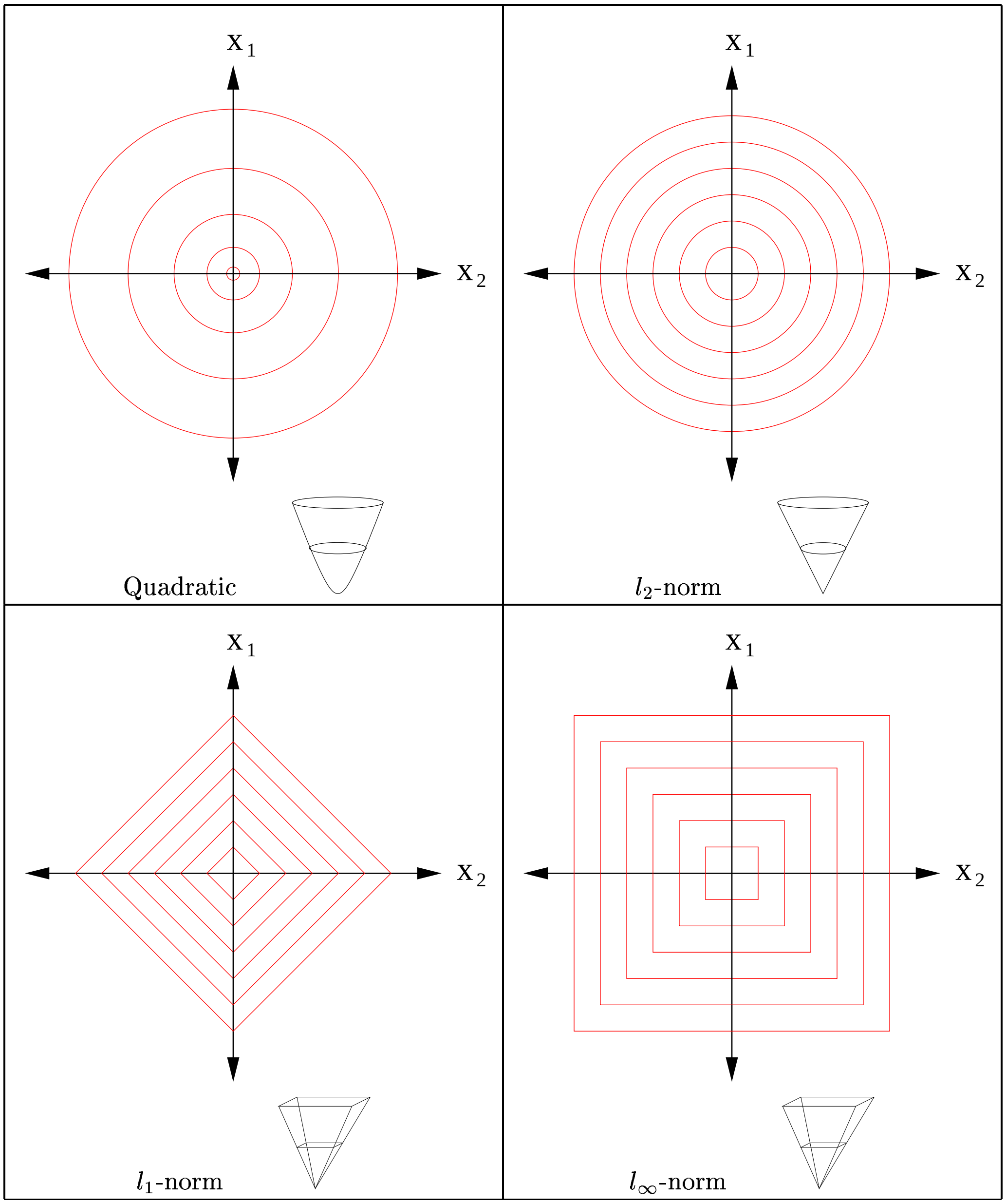


Figure 7.1: Geometric interpretation of cost functions

The basic strategy is to consider only a finite number of decision variables, so that the infinite horizon problem reduces to a a finite-dimension mathematical program.

We transform the infinite horizon problem to a finite horizon problem with a terminal state penalty by considering the free evolution of only the stable modes on the infinite horizon. We obtain the transformation using the following terminal penalty

$$V(x) = \sum_{k=0}^{\infty} \|\bar{Q}CA_s^k x\|_p,$$

where A_s is the restriction of A to the stable subspace of A. For the majority of systems an analytic expression for $V(x)$ is unavailable. One simple strategy to generate a stable approximation for $V(x)$ is to assume that the nonzero eigenvalues of A_s are nondefective. This assumption allows us to upper bound the sum with a Lyapunov function. Consider the Jordan decomposition

$$A_s = S \begin{bmatrix} \Lambda & 0 \\ 0 & J_{n_0}(0) \end{bmatrix} S^{-1},$$

where the diagonal matrix Λ contains the nonzero eigenvalues of A_s and n_0 is the algebraic multiplicity of the zero eigenvalue. Because the Jordan block $J_{n_0}(0)$ is nilpotent, we have

$$A_s^n = \begin{bmatrix} S_\Lambda & S_0 \end{bmatrix} \begin{bmatrix} \Lambda^n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} (S^{-1})_\Lambda \\ (S^{-1})_0 \end{bmatrix}, \tag{7.4a}$$

$$= S_\Lambda \Lambda^n (S^{-1})_\Lambda, \tag{7.4b}$$

for all $n \geq n_0$. If we consider the coordinate transformation $z = (S^{-1})_\Lambda x$, we generate the following upper bound

$$\bar{V}(x) = \sum_{k=0}^{n_0-1} \|\bar{Q}CA^k x\|_p + \theta^T |z|,$$

where $|z|$ is a vector whose entries are the absolute value of the associated entries of z and

$$\theta^{(j)} = \sum_{k=n_0}^{\infty} \|\bar{Q}CS_\Lambda \Lambda^k e_j\|_p,$$

where the vector e_j is the unit vector whose j^{th} entry is 1. The validity of this bound follows directly from the subadditivity of norms:

$$\sum_{k=n_0}^{\infty} \|\bar{Q}CA_s^k x\|_p =$$

$$\sum_{k=n_0}^{\infty} \|\bar{Q}CS_\Lambda \Lambda^k (S^{-1})_\Lambda x\|_p =$$

$$\sum_{k=n_0}^{\infty} \|\bar{Q}CS_\Lambda \Lambda^k (z^{(1)} e_1 + \ldots + z^{(n-n_0)} e_{n-n_0})\|_p$$

$$\leq \sum_{k=n_0}^{\infty} \sum_{j=1}^{n-n_0} |z^{(j)}| \, \|\bar{Q}CS_\Lambda \Lambda^k e_j\|_p = \theta^T |z|.$$

Lemma 7.2.1 $\bar{V}(x) \geq \|QCx\|_p + \bar{V}(A_s x)$.

Proof. If follows from (7.4a) that $(S^{-1})_\Lambda A_s x = \Lambda z$. Hence, we have

$$\begin{aligned}
&\bar{V}(A_s x) - \bar{V}(x) = \\
&\qquad \|QCS_\Lambda \Lambda^{n_0} z\|_p + \theta^T|\Lambda z\| - \left(\|QCx\|_p + \theta^T|z|\right).
\end{aligned}$$

Expanding $\theta^T|z|$, we generate the follow inequality.

$$\begin{aligned}
\theta^T|x| &= \sum_{j=1}^{n-n_0} |z^j|\, \|\bar{Q}CS_\Lambda \Lambda_0^n e_j\|_p + \\
&\quad |z^j| \sum_{k=n_0}^{\infty} \|\bar{Q}CS_\Lambda \Lambda^k \Lambda e_j\|_p, \\
&\geq \|\bar{Q}CS_\Lambda \Lambda^{n_0} z\|_p + \\
&\quad \sum_{k=n_0}^{\infty} \sum_{j=1}^{n-n_0} |\Lambda^{(j)} z^{(j)}|\, \|\bar{Q}CS_\Lambda \Lambda^k e_j\|_p, \\
&= \|\bar{Q}CS_\Lambda \Lambda^{n_0} z\|_p + \theta^T|\Lambda z|.
\end{aligned}$$

Hence the lemma follows. □

We formulate the finite-horizon regulator as the solution to

$$\begin{aligned}
&\Phi_N^*(\hat{x}_k) = \\
&\qquad \min_{u_k, x_k} \sum_{k=0}^{N-1} \|\bar{R}u_k\|_p + \|\bar{Q}y_k\|_p + \bar{V}(x_N),
\end{aligned} \tag{7.5}$$

subject to (7.1), the initial condition $x_0 = \hat{x}_j$, (7.3), and

$$F^T x_N = 0, \tag{7.6}$$

where the columns of F span the orthogonal complement of the stable subspace of A. An ordered Schur decomposition of A yields an orthogonal representation of F. In the absence of the constraints (7.3), choosing $N \geq n$ is sufficient to guarantee feasibility. With the presence of inequality constraints, feasibility is obtained for stable systems if N is sufficiently large such that

$$x_N \in \mathcal{O}_\infty, \tag{7.7}$$

where the set $\mathcal{O}_\infty$ is positive invariant and contained within the feasible region specified by (7.3). Details concerning the properties and construction of $\mathcal{O}_\infty$ are available in (Gilbert and Tan 1991). For unstable systems, we also require that the state $\hat{x}_j$ is contained in the set of constrained stabilizable states and N is sufficiently large such that (7.6) is feasible.

Proposition 7.2.2 *If a feasible solution exists, then the origin is an asymptotically stable fixed point for the closed-loop system.*

Proof. Stability follows from the continuity of $\Phi_N^*(\cdot)$. To demonstrate convergence, let

$$\{u_{k|k}, \ldots, u_{k+N-1|k}\}$$

denote the minimizing sequence at time index k. Since the sequence

$$\{u_{k+1|k}, \ldots, u_{k+N|k}, 0\}$$

is also admissible at time $k+1$, we have from Lemma 7.2.1 that

$$\Phi_N^*(\hat{x}_k) - \Phi_N(\hat{x}_{k+1|k}) \geq \left(\|\bar{R}u_{k|k}\|_p + \|\bar{Q}y_{k|k}\|_p\right).$$

The sequence $\{\Phi_N^*(x_k)\}_{k=0}^{\infty}$ is convergent, because it is nonincreasing and bounded below. Hence,

$$\left(\|\bar{R}u_{k|k}\|_p + \|\bar{Q}y_{k|k}\|_p\right) \to 0$$

as $k \to \infty$. Because (C, A) is detectable, we have $x_k \to 0$. Therefore, the regulator is asymptotically stable as claimed. □

7.2.1 Linear Programming Formulations

With either an l_1 or l_∞ criterion, we may transform the optimal control problem to a linear program by introducing auxiliary variables. We formulate (7.5) with an l_1 criterion as the following linear program

$$\Phi_N^*(\hat{x}_j) = \min_{x_k, u_k, \rho_k, \eta_k, \gamma_k, z_N} \sum_{k=0}^{N-1} e^T\rho_k + e^T\eta_k + \sum_{k=0}^{n_0-1} e^T\gamma_k + \theta^T z_N,$$

subject to (7.1), the initial condition $x_0 = \hat{x}_j$, (7.3), and (7.6), where the non-negative vectors ρ_k, η_k, γ_k, and z_N are specified by the following linear inequalities

$$-\rho_k \leq \bar{R}u_k \leq \rho_k, \qquad -\eta_k \leq \bar{Q}Cx_k \leq \eta_k,$$
$$-\gamma_k \leq \bar{Q}CA_s^k x_N \leq \gamma_k, \; -z_N \leq (S^{-1})_\Lambda x_N \leq z_N.$$

With an l_∞ criterion, we formulate (7.5) as the following linear program

$$\Phi_N^*(\hat{x}_j) = \min_{x_k, u_k, \rho_k, \eta_k, \gamma_k, z_N} \sum_{k=0}^{\infty} \rho_k + \eta_k + \sum_{k=0}^{n_0-1} \gamma_k + \theta^T z_N,$$

subject to (7.1), the initial condition $x_0 = \hat{x}_j$, (7.3), and (7.6), where the non-negative scalars ρ_k ,η_k, and γ_k and the vector z_N are specified by the following linear inequalities

$$-\rho_k e \leq \bar{R}u_k \leq \rho_k e, \;\; -\eta_k e \leq \bar{Q}Cx_k \leq \eta_k e,$$
$$-\gamma_k e \leq \bar{Q}CA_s^k x_N \leq \gamma_k e,$$
$$-z_N \leq (S^{-1})_\Lambda x_N \leq z_N.$$

The variable e is the vector of ones.

7.3 MPC with an l_1 Norm Objective

Consider the regulation of the following non-minimum phase system

$$y(s) = \frac{s-3}{3s^2+4s+2}u(s),$$

sampled at frequency of 10 Hertz with an initial state disturbance of $x_0 = [1,\ 1]^T$. A horizon length of $N = 30$ was chosen for both examples. For simplicity we ignore inequality constraints, because they add little to the theme of the discussion on qualitative performance. Figure 7.2 shows the comparison of the closed-loop responses between an l_1 criterion and a quadratic criterion with tuning parameters $\bar{Q} = 5$ and $\bar{R} = 1$. The simulation indicates the l_1 formulation forces the state to the origin in finite time as opposed to the quadratic programming formulation, where the state exponentially approaches the origin. Further simulations indicate that the dead-beat policy holds for all initial conditions. The finite horizon problem is also equivalent to the infinite horizon problem, because the l_1 formulation forces the state to the origin in finite time. Forcing the state to the origin in finite time is appealing for servo regulation. However, dead-beat control may yield poor closed-loop performance in process control applications. The poor performance becomes evident when state noise is added to the simulation. Figure 7.3 shows a comparison of closed-loop responses when state noise is added. The deviation from the target is less for the l_1 formulation; at the same time the dead-beat performance causes aggressive control action. In many situations this high-gain control is undesirable.

In addition to yielding dead-beat performance, the l_1 formulation results in idle control performance when the input penalty $\bar{R}$ is large relative to $\bar{Q}$. Figure 7.4 shows the comparison of the closed-loop responses between the l_1 criterion and the quadratic criterion with tuning parameters $\bar{Q} = 1$ and $\bar{R} = 5$. The simulation indicates that the optimal policy for the l_1 formulation is no control action. For the given tuning, the idle policy holds regardless of the initial conditions and the horizon length. Although the qualitative performance, e.g. the settling time, between the quadratic and l_1 criterion is not appreciably different for the example, idle control defeats the purpose of implementing a control system. The reason for the similarity between the open-loop response (idle control policy) and the closed loop response (quadratic criterion) is that the large input penalty $\bar{R}$ relative to the output penalty $\bar{Q}$ pacifies the controller and, therefore, does not place the closed-loop poles far from the open-loop poles. So, unlike the dead-beat policy, with the addition of disturbances, the qualitative response of the idle policy is similar to response of the quadratic formulation.

The two examples demonstrates that the l_1 formulation yields different qualitative responses depending on the selection of the tuning parameters. This dichotomy is in direct contrast to the quadratic formulation, where the qualitative response is the same regardless of the tuning; i.e. the qualitative response is always exponential convergence. The difference between the two formulations is analogous to the difference between a positive definite quadratic program and a linear program. Whereas the solution to the quadratic program may reside in the interior of the feasible region, the solution to a linear program always resides at an extreme point of the feasible region.

We can specifically attribute the differences between the l_1 and quadratic formulation to the nonsmoothness of the objective function. We can interpret the input and output stage costs as competing exact penalties, because the objective function for the l_1 criterion is a sum of norms (an introductory explanation of exact penalties may be found in Fletcher (Fletcher 1987)). The purpose of exact penalties is to recast the constrained optimization

$$\min_x \{f(x) \ : \ g(x) = 0\}$$

as the equivalent unconstrained optimization

$$\min_x f(x) + \lambda \|g(x)\|_p.$$

If $\lambda > 0$ is sufficiently large (greater than the dual norm of the Lagrange multiplier associated with the constraint $g(x) = 0$), then the solutions to the two optimization problems are equivalent. Hence, we may view the terms $\|\bar{R}u_k\|_1$ as penalties for the constraint $u_k = 0$ and the terms $\|\bar{Q}y_k\|_1$ as penalties for the constraint $y_k = 0$. When the input penalty $\bar{R}$ is sufficiently large, the exact penalty $\|\bar{R}u_k\|_1$

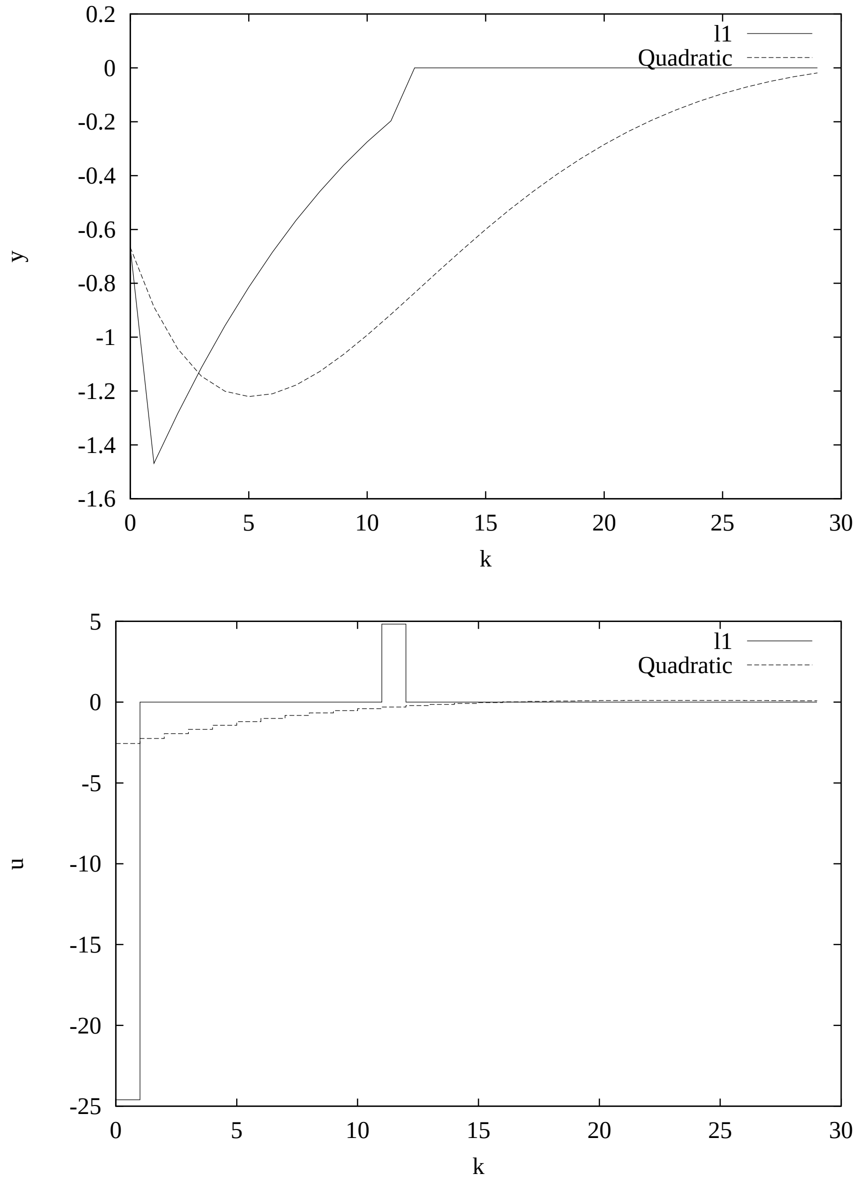
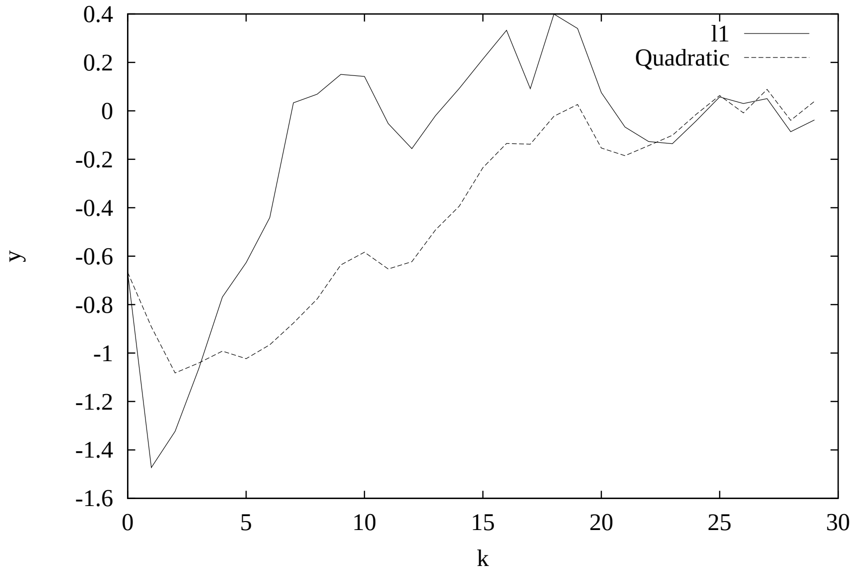


Figure 7.2: Comparison of input and output responses for $Q = 5$ and $R = 1$

becomes the binding constraint $u_k = 0$. Likewise, when the output penalty $\bar{Q}$ is sufficiently large, the exact penalty $\|\bar{Q}y_k\|_1$ becomes the binding constraint $y_k = 0$. In particular, the two penalties compete respectively for dead-beat and idle control performance. We also expect the same qualitative behavior for the l_∞ formulation, because the nonsmoothness is present for any l_p formulation,

We demonstrate the effect of the nonsmoothness of the objective function geometrically with a simple scalar example. Consider the following single stage optimal control problem

$$\min_{u_0} \Theta = |x_1| + r|u_0|,$$



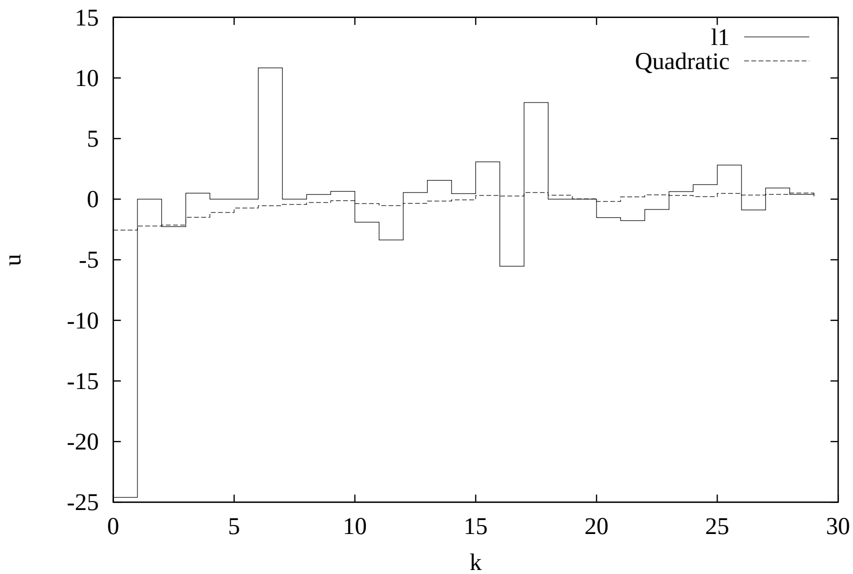


Figure 7.3: Comparison of input and output responses for $Q = 5$ and $R = 1$ with state noise

subject to the scalar system

$$x_1 = ax_0 + bu_0.$$

Recognize that because both the state and input are scalar, this example encompasses all l_p norm formulations. Figure 7.5 shows the graph of Θ as a function of u_0. It is evident from the graph that if $r > b$, then the optimal solution is $u_0 = 0$, because the slope of the middle section is negative. Likewise, if $r < b$, the optimal solution is $u_0 = -\frac{ax_0}{b}$, which yields dead-beat control, because the slope of the middle section is positive. If $r = b$, the optimal solution is not unique. Both solutions are optimal,

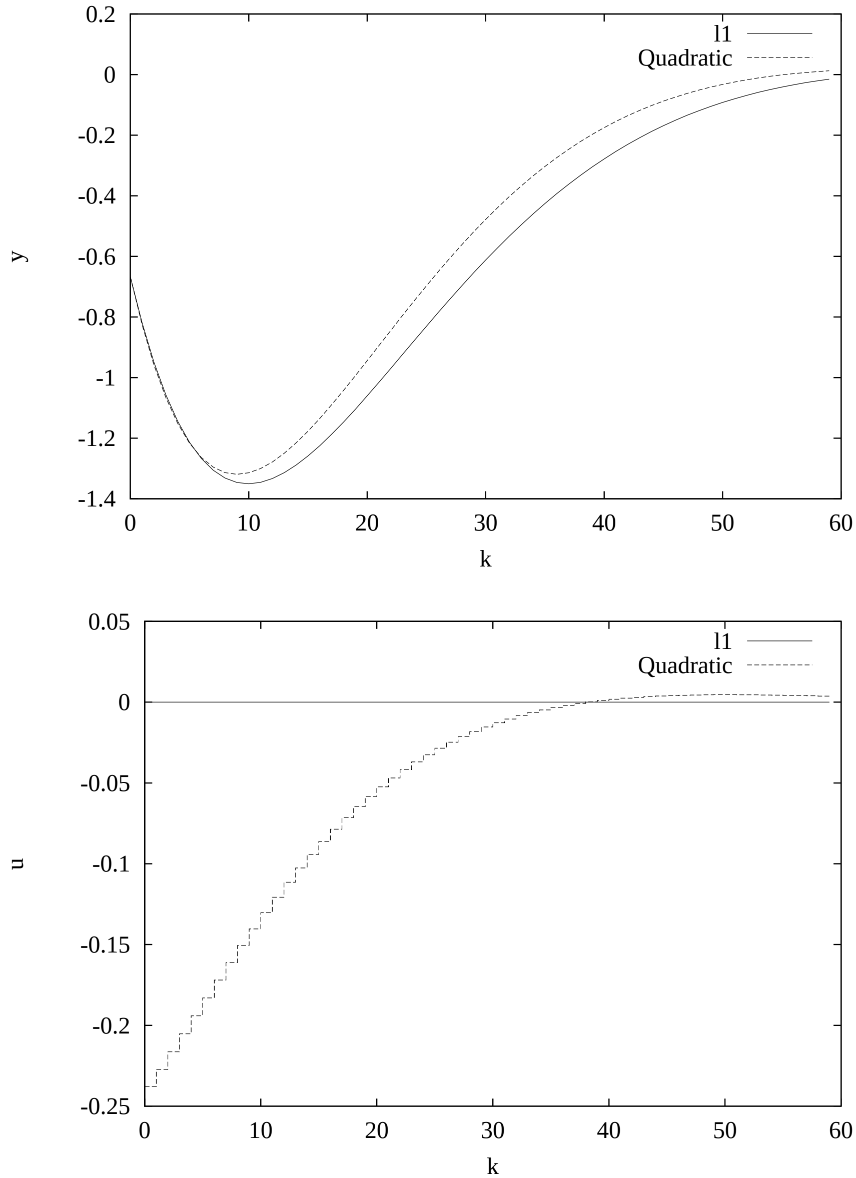


Figure 7.4: Comparison of input and output responses for $Q = 1$ and $R = 5$

including all solutions in between: either $0 \leq u_0 \leq -ax_0/b$ or $-ax_0/b \leq u_0 \leq 0$. If we consider the quadratic criterion

$$\Theta' = x_1^2 + ru_0^2,$$

then the optimal solution is

$$u_0 = -\frac{ab}{b^2 + r} x_0.$$

In contrast to the l_1 formulation, the quadratic control is neither dead-beat nor idle for $r > 0$.

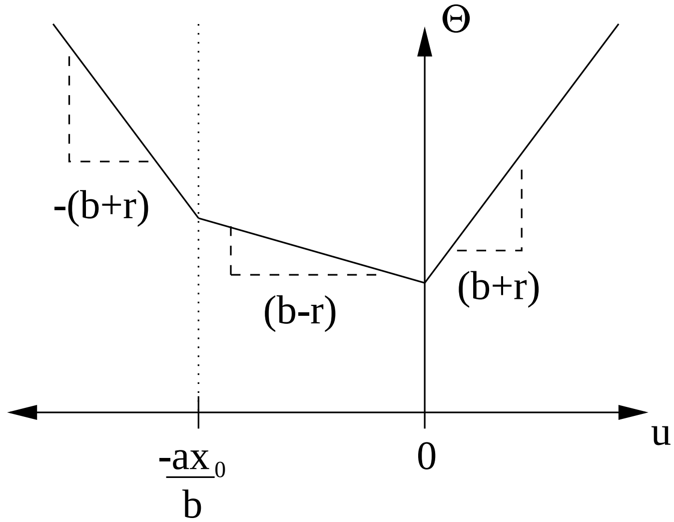


Figure 7.5: Graph and slopes of the cost function Θ.

At this stage we are confronted with the question as to whether the l_p formulation is preferable to a quadratic formulation. In addition to the numerical advantages offered by linear programs, for many applications the actual control specifications translate more naturally into an l_p criterion than a quadratic criterion. However, the sensitivity of closed-loop behavior for the l_1 formulation is disconcerting, because the tuning parameters must be chosen judiciously to exclude undesirable performance. Not only does one have to be wary of the implications of dead-beat or idle performance, non-uniqueness of the control presents potential problems, because erratic closed-loop behavior may result. We expect that additional measures such as input velocity penalties would help counteract the aggressive control behavior. However, the additional measures would only compensate for and not alter the fundamental behavior of l_p formulations.

7.4 Duality and Hahn-Banach

To provide greater insight in the problem, we can examine the dual optimization problem. Since the l_1 problem is non-differentiable, we require the abstract setting of a Banach space, in particular the Hahn-Banach theorem. Keerthi (1986) also analyzed the qualitative performance of the l_1 formulation using dynamic programming arguments. His insightful work provided the impetus for this study.

Theorem 7.4.1 (Hahn-Banach) *Let $D \in \mathbb{R}^{n\times m}$, then*

$$\min_{Dx=d} \|x\| = \max_{\|D^T y\|^* \leq 1} \langle d, y \rangle, \tag{7.8}$$

where

$$\|x\|^* := \max_{\|y\|\leq 1} \langle x, y \rangle.$$

Furthermore, the optimal solution x^* is aligned with the optimal $D^T y^*$, i.e.

$$\langle x^*, D^T y^* \rangle = \|x^*\| \|D^T y^*\|^*.$$

Proof. See Luenberger (1969) for details. $\square$

A direct application of Theorem 7.4.1 is the following proposition.

Proposition 7.4.2 *Consider the following optimal control problem*

$$\Phi_N^*(x_0) := \min\left\{\sum_{k=0}^{N-1} \|\bar{R}u_k\|_1 + \|\bar{Q}y_{k+1}\|_1 \;:\; \Sigma(x_0)\right\},$$

then the dual problem is

$$\max \sum_{k=1}^{N} \langle QCA^k x_0, \lambda_k\rangle \tag{7.9}$$

subject to the constraints

$$\left\|\textstyle\sum_{j=k}^{N}\langle QCA^{j-k}BR^{-1}, \lambda_j\rangle\right\|_\infty \le 1, \quad \text{for } k = 1, \dots N; \tag{7.10a}$$

$$\|\lambda_k\|_\infty \le 1, \;\; k = 1, \dots, N. \tag{7.10b}$$

Furthermore, the optimal solutions are aligned:

$$\Phi^*(x_0) = \sum_{k=1}^{N}\langle u_{k-1}, \sum_{j=k}^{N}\langle QCA^{j-k}BR^{-1}, \lambda_j\rangle\rangle - \sum_{k=1}^{N}\langle x_k, \lambda_k\rangle.$$

Proof. If we consider the change of variables

$$u_k \leftarrow \bar{R}u_k, \qquad y_k \leftarrow \bar{Q}y_k,$$

then

$$y_k - \sum_{j=0}^{k-1} QCA^j R^{-1} u_j = \bar{Q}CA_k x_0,$$

for $k = 1, \dots N$. Because l_∞ is the dual of l_1, the proposition follows directly from Theorem 7.4.1. □

An immediate consequence of Proposition 7.4.2 is the following result.

Proposition 7.4.3 *If*

$$\max_{1\le i\le m} \sum_{j=0}^{N-1} \|QCA^jBR_i^{-1}\|_1 < 1,$$

where R_i^{-1} is the i^{th} column of R^{-1}, then $\{u_0 = 0\} \in \arg\min \Phi(x_0)$ for all $x_0 \in \mathbb{R}^n$.

Proof. By the subadditivity of norms and the identity $\|A^T\|_\infty = \|A\|_1$,

$$\left\|\sum_{j=1}^{N}\langle QCA^{j-1}BR^{-1}, \lambda_j\rangle\right\|_\infty \le \max_{1\le i\le m}\sum_{j=1}^{N} \|(QCA^{j-1}BR_i^{-1}\|_1\|\lambda_k\|_\infty \le \max_{1\le i\le m}\sum_{j=0}^{N-1}\|QCA^jBR_i^{-1}\|_1 < 1.$$

Because of the alignment property, $u_0 \neq 0$ if and only if

$$\left\| \sum_{j=k}^{N} \langle QCA^{j-k}BR^{-1}, \lambda_j \rangle \right\|_\infty = 1.$$

Hence the proposition follows. □

For the example with $\bar{Q} = 1$ and $\bar{R} = 5$,

$$\sum_{j=0}^{\infty} \|QCA^j BR^{-1}\|_1 \approx 0.325.$$

Conjecture 7.4.4 *Assume (A, B) is controllable and (A, C) observable. If*

$$\min_{1 \leq i \leq m} \sum_{j=0}^{\infty} \|QCA^j BR_i^{-1}\|_1 > 1,$$

then there exist $N^* < \infty$ such that $\Phi_N^* = \Phi_\infty^*$.

Sketch of a possible proof: For $k = 1, \ldots n$, we have the following equality

$$\begin{gathered} \mathcal{A}_{ki} := \sum_{j=k}^{\infty} \langle QCA^{j-k}BR_i^{-1}, \lambda_j \rangle = \\ \sum_{j=0}^{n-k-1} \langle QCA^j BR_i^{-1}, \lambda_{j+k} \rangle + \\ \sum_{j=0}^{\infty} \langle A^{n-k}BR_i^{-1}, A^j C^T Q \lambda_{j+n} \rangle \end{gathered}$$

We can then represent the objective function as follows.

$$\begin{aligned} \Phi_\infty^* &= \max_{\lambda_k} \sum_{k=1}^{\infty} \langle QCA^k x_0, \lambda_k \rangle, \\ &= \sum_{k=1}^{n-1} \langle QCA^k x_0, \lambda_k \rangle + \sum_{k=n}^{\infty} \langle QCA^k x_0, \lambda_k \rangle, \\ &= \sum_{k=1}^{n-1} \langle QCA^k x_0, \lambda_k \rangle + \\ & \quad \langle A^n x_0, \sum_{j=0}^{\infty} A^{jT} C^T Q \lambda_{j+n} \rangle. \end{aligned}$$

Because Σ is controllable, we have

$$A^n x_0 \in \operatorname{ran}[BR^{-1}, ABR^{-1}, \ldots, A^{n-1}BR^{-1}].$$

Hence we have

$$\begin{aligned}\Phi_\infty^* &= \sum_{k=1}^{n-1}\langle QCA^k x_0, \lambda_k\rangle + \\ &\quad \sum_{v=1}^{n}\sum_{i=1}^{m}\alpha_{vi}\langle A^{n-v}BR_i^{-1}, \sum_{j=0}^{\infty}A^{jT}C^TQ\lambda_{j+n}\rangle, \\ &= \sum_{k=1}^{n-1}\langle QCA^k x_0, \lambda_k\rangle - \\ &\quad \sum_{k=1}^{n}\sum_{i=1}^{m}\alpha_{ki}\sum_{j=0}^{n-k-1}\langle QCA^jBR^{-1}, \lambda_{j+k}\rangle + \alpha_{ki}\mathcal{A}_{ki}, \\ &= \sum_{k=1}^{n-1}\langle \beta_k, \lambda_k\rangle + \sum_{k=1}^{n}\sum_{i=1}^{m}\alpha_{ki}\mathcal{A}_{ki}.\end{aligned}$$

Suppose that $\|\lambda_k\|_\infty = 1$ for all $k \in \mathbb{N}$. By assumption, the sequence $\{\lambda_k\}$ cannot be extremal with respect to $\mathcal{A}_{ki}$. Otherwise the constraint $\|\mathcal{A}_{ki}\|_\infty \leq 1$ is violated.

$$\Phi_\infty^* \leq \sum_{k=1}^{n-1}\langle \beta_k, \lambda_k\rangle + \sum_{k=1}^{n}\sum_{i=1}^{m}|\alpha_{ki}|.$$

Somehow, the heart of the conjecture, this implies that there exists an integer N^* such that $\|\lambda_k\|_\infty < 1$ for $k \geq N^*$. The remaining steps follow from the observability assumption and the alignment property . $\diamond$

Let us consider again the single stage optimal control problem

$$\min_{u_0} \|x_1\| + r\|u_0\|$$

subject to the scalar system

$$x_1 = ax_0 + bu_0.$$

Using Proposition 7.4.2, we construct the dual problem

$$\max_{\lambda} x_0\lambda$$

subject to

$$|b\lambda| \leq r, \qquad |\lambda| \leq 1.$$

The solution to both the primal and dual problems satisfy the following alignment or complementarity condition

$$u_0(b\lambda) + x_1\lambda = |x_1| + r|u_0|. \tag{7.11}$$

By inspection we note $x_1 \neq 0$ if and only if $|\lambda| = 1$. Otherwise the inequality (7.11) cannot be satisfied: $|b\lambda| \leq 1$. Likewise, $u_0 \neq 0$ if and only if $|b\lambda| = r$. These conditions are depicted in Figure 7.6. The preceding arguments are simply extensions of these elementary results.

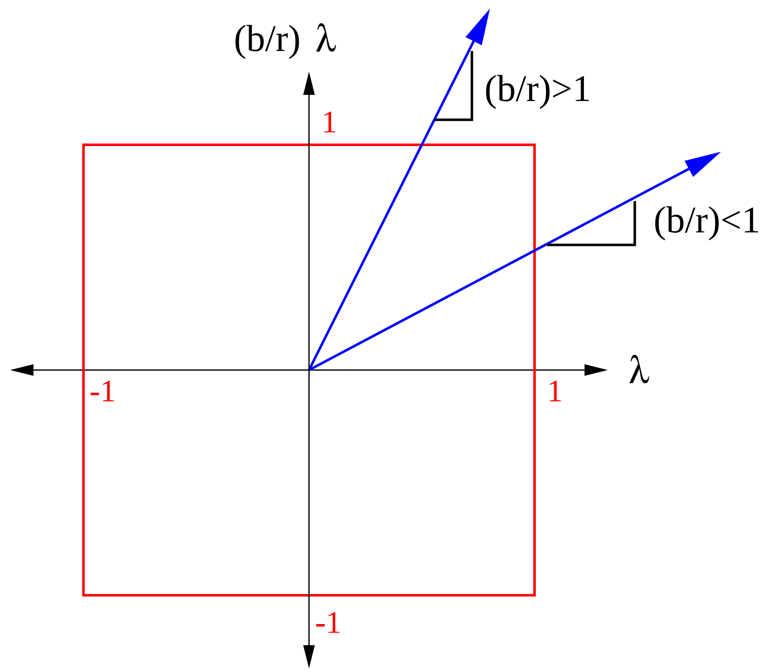


Figure 7.6: The duality conditions.

7.5 Conclusion

The main contribution of this chapter has been to illustrate some of the consequences of using MPC with an l_p criterion. Our motivation for studying the l_p criterion was that for both l_1 and l_∞ criterion the resulting optimization can be formulated as a linear program. Linear programming formulations are desirable, because they are computationally less demanding than standard quadratic programming formulations. Furthermore, theoretical issues such as stability are a straightforward extension of the results available for the quadratic criterion. However, performance issues raise questions concerning the suitability of the l_p criterion for MPC. Although possessing desirable theoretical and numerical properties, l_p formulations suffer many practical drawbacks. The main consequence of the l_p criterion is that it may yield either dead-beat or idle control response. Both of these types of responses may be unsuitable for process control application.

While our arguments have been mostly qualitative, it is evident that the culprit is the nonsmoothness of the objective function. The nonsmoothness causes the stage cost functions to act as competing exact penalties for the constraints $u = 0$ and $y = 0$. For the scalar system, the behavior is simple to understand. Extending these results to higher dimension systems is more difficult and is currently unresolved as illustrated in Section 7.4.

Chapter 8

Application of Interior-Point Methods to Model Predictive Control [1]

The MPC methodology is appealing to the practitioner because input and state constraints can be explicitly accounted for in the controller. A practical disadvantage is its computational cost, which has tended to limit MPC applications to linear processes with relatively slow dynamics. For such problems, the optimal control problem to be solved at each stage of MPC is a convex quadratic program. While robust and efficient software exists for the solution of unstructured convex quadratic programs, significant improvements often can be made by exploiting the structure of the MPC subproblem.

When input and state constraints are not present, MPC with an infinite horizon is simply the well-known linear-quadratic regulator problem. Even when constraints are present, the infinite-horizon MPC problem generally reduces to a linear-quadratic regulator after a certain number of stages (c.f. (Chmielewski and Manousiouthakis 1996, Scokaert and Rawlings 1998, Sznaier and Damborg 1987)) and therefore can be recast as a finite-dimensional quadratic program. Since this quadratic program can be large, with many stages, it is important that algorithms be efficient for problems with long horizons.

Unconstrained discrete-time linear-quadratic optimal control problems can be solved by using a discrete-time Riccati equation. The computational cost of this algorithm is linear in the horizon length N. A different formulation obtained by eliminating the state variables results in an unconstrained quadratic function whose Hessian is dense, with dimensions that grow linearly in N. The cost of minimizing this quadratic function is cubic in N, making it uncompetitive with the Riccati approach in general. There is a third option, however—an optimization formulation in which the states are retained explicitly as unknowns in the optimization and the model equation is retained as a constraint. The optimality conditions for this formulation reveal that the adjoint variables are simply the Lagrange multipliers for the model equation and that the problem can be solved by factoring a matrix whose dimension again grows linearly with N. In this formulation, however, the matrix is banded, with a bandwidth independent of N, so the cost of the factorization is linear rather than cubic in N. The discrete-time Riccati equation can be interpreted as a block factorization scheme applied to this matrix.

Traditionally, the discrete-time Riccati equation is obtained by using dynamic programming to solve the unconstrained linear optimal control problem. The essential idea in dynamic programming is to work stage-by-stage through the problem in reverse order, starting with the final stage N. The optimization problem reduces to a simpler problem at each stage. (See Bertsekas (1987) for further details.) Block factorization, like dynamic programming, exploits the multi-staged nature of the optimization

[1] **Portions of this chapter were published in Rao et al. (1998) and Rao, Campbell, Rawlings and Wright (1997)**

problem. The key difference is that the block factorization approach tackles the problem explicitly, whereas dynamic programming tackles the problem semi-implicitly by using Bellman's principle of optimality. The explicit treatment allows greater flexibility, however, since the block factorization approach retains its inherent structure even when inequality constraints are added to the formulation.

When constraints are present, the scheme for unconstrained problems must be embedded in an algorithmic framework that determines which of the inequalities are active and which are inactive at the optimum. At each iteration of the outer algorithm, however, the main computational operation is the solution of a set of linear equations whose structure is very like that encountered in the unconstrained problem. Hence, the cost of performing each iteration of the outer algorithm is linear in the number of stages N. This observation has been made by numerous authors, in the context of outer algorithms based on both active-set and interior-point methods. Glad and Jonson (1984) and Arnold, Tatjewski and Wolochowicz (1994) demonstrate that the factorization of a structured Lagrangian in an optimal control problem with a Bolza objective for an active set framework yields a Riccati recursion. Wright (1993, 1997*a*), Steinbach (1994), and Lim, Moore and Faybusovich (1996) investigate the Bolza control problem in an interior-point framework.

In this chapter we present an MPC algorithm based on an interior-point method, in which a block factorization is used at each iteration to obtain the search direction for the interior-point method. Our work differs from earlier contributions in that the formulation of the optimal control problem is tailored to the MPC application, the interior-point algorithm is based on Mehrotra's algorithm (Mehrotra 1992) (whose practical efficiency on general linear and quadratic programming problems is well documented), and the linear system at each interior-point iteration is solved efficiently by a Riccati recursion. We compare our approach with the alternative of using the model equation to eliminate the states, yielding a dense quadratic program in the input variables alone, and present results obtained for three large industrial problems.

We use order notation in the following (standard) way: If a matrix, vector, or scalar quantity M is a function of another matrix, vector, or scalar quantity E, we write $M = O(\|E\|)$ if there is a constant β such that $\|M\| \leq \beta\|E\|$ for all $\|E\|$ sufficiently small. We write $M = \Theta(\|E\|)$ if there is a constant β such that $\|E\|/\beta \leq \|M\| \leq \beta\|E\|$.

We say that a matrix is "positive diagonal" if it is diagonal with positive diagonal elements. The term "nonnegative diagonal" is defined correspondingly. We use SPD as an abbreviation for "symmetric positive definite" and SPSD as an abbreviation for "symmetric positive semidefinite."

8.1 Problem Statement

In this chapter we work with a general form of the MPC problem, which contains all the features discussed in Chapter 6: finite horizon, endpoint constraints, and soft constraints. This general form is

$$\min_{u,x,\epsilon} \Phi(u,x,\epsilon) = \sum_{k=0}^{N-1} \frac{1}{2}(x_k^T Q x_k + u_k^T R u_k + 2x_k^T M u_k + \epsilon_k^T Z \epsilon_k) + z^T \epsilon_k + x_N^T \bar{Q}_N x_N, \tag{8.1}$$

subject to

$$\begin{aligned}
x_0 &= \hat{x}_j, \quad \text{(fixed)} && (8.2a)\\
x_{k+1} &= Ax_k + Bu_k, \quad k = 0,1,\ldots,N-1, && (8.2b)\\
Du_k - Gx_k &\leq d, \quad k = 0,1,\ldots,N-1, && (8.2c)\\
Hx_k - \epsilon_k &\leq h, \quad k = 1,2,\ldots,N, && (8.2d)\\
\epsilon_k &\geq 0, \quad k = 1,2,\ldots,N, && (8.2e)\\
Fx_N &= 0. && (8.2f)
\end{aligned}$$

We assume throughout that the matrices in (8.1) satisfy the properties

$$R \text{ is PSD}, \qquad Z \text{ is SPSD}, \qquad \begin{bmatrix} Q & M \\ M^T & R \end{bmatrix} \text{ is SPSD.} \tag{8.3}$$

Note that the last property holds for the matrices considered in Section 6.3, since in making the substitutions to obtain the form (6.14) we obtain

$$\begin{aligned} \begin{bmatrix} Q & M \\ M^T & R \end{bmatrix} &\leftarrow \begin{bmatrix} Q + L^T(R+S)L & -L^TS & L^T(R+S) \\ -SL & S & -S \\ (R+S)L & -S & (R+S) \end{bmatrix} \\ &= \begin{bmatrix} Q & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} L^T \\ -I \\ I \end{bmatrix} S \begin{bmatrix} L & -I & I \end{bmatrix} + \begin{bmatrix} L^T \\ 0 \\ I \end{bmatrix} R \begin{bmatrix} L & 0 & I \end{bmatrix}, \end{aligned}$$

which is a sum of SPSD matrices and is therefore itself SPSD.

8.2 The Interior-Point Method

In this section, we describe our interior-point-based approach for solving the MPC problem (8.1), (8.2). We start with a general description of the interior-point method of choice for linear and convex quadratic programming: Mehrotra's predictor-corrector algorithm. The remaining sections describe the specialization of this approach to MPC, including the use of the Riccati approach to solve the linear subproblem, handling of endpoint constraints, and hot starting.

8.2.1 Mehrotra's Predictor-Corrector Algorithm

Active set methods have proved to be efficient for solving quadratic programs with general constraints. The interior-point approach has proved to be an attractive alternative when the problems are large and convex. In addition, this approach has the advantage that the system of linear equations to be solved at each iterate has the same dimension and structure throughout the algorithm, making it possible to exploit any structure inherent in the problem. The most widely used interior-point algorithms do not require a feasible starting point to be specified. In fact, they usually generate infeasible iterates, attaining feasibility only in the limit. From a theoretical viewpoint, interior-point methods exhibit polynomial complexity, in contrast to the exponential complexity of active-set approaches.

In this section, we sketch an interior-point method for general convex quadratic programming problems and discuss its application to the specific problem (8.1). A more complete description is given by Wright (1997*b*).

Consider the following convex quadratic program

$$\min_w \Phi(w) = \frac{1}{2} w^T Q w + c^T w, \text{ subject to } Fw = g,\ Cw \le d, \tag{8.4}$$

where Q is an SPSD matrix. The Karush-Kuhn-Tucker (KKT) conditions for optimality are that there exist vectors π^* and λ^* such that the following conditions are satisfied for $(w, \pi, \lambda) = (w^*, \pi^*, \lambda^*)$:

$$\begin{aligned} Qw + F^T\pi + C^T\lambda + c &= 0, \\ -Fw + g &= 0, \\ -Cw + d &\ge 0, \\ \lambda &\ge 0, \\ \lambda_j(-Cw + d)_j &= 0, \qquad j = 1, 2, \ldots, m, \end{aligned}$$

where m is the number of rows in the matrix C. Because the objective function is convex, the KKT conditions are both necessary and sufficient for optimality. By introducing a vector t of slacks for the constraint $Cw \leq d$, we can rewrite these conditions in a slightly more convenient form:

$$\mathcal{F}(w, \pi, \lambda, t) = \begin{bmatrix} Qw + F^T\pi + C^T\lambda + c \\ -Fw + g \\ -Cw - t + d \\ T\Lambda e \end{bmatrix} = 0, \tag{8.5a}$$

$$(\lambda, t) \geq 0, \tag{8.5b}$$

where T and Λ are diagonal matrices defined by

$$T = \operatorname{diag}(t_1, t_2, \ldots, t_m), \qquad \Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m),$$

and $e = (1, 1, \ldots, 1)^T$.

Primal-dual interior-point methods generate iterates $(w^i, \pi^i, \lambda^i, t^i)$, $i = 1, 2, \ldots$, with $(\lambda^i, t^i) > 0$ that approach feasibility with respect to the conditions (8.5a) as $i \to \infty$. The search directions are Newton-like directions for the equality conditions in (8.5a). Dropping the superscript and denoting the current iterate by (w, π, λ, t), we can write the general linear system to be solved for the search direction as

$$\begin{bmatrix} Q & F^T & C^T & \\ -F & & & \\ -C & & & -I \\ & & T & \Lambda \end{bmatrix} \begin{bmatrix} \Delta w \\ \Delta \pi \\ \Delta \lambda \\ \Delta t \end{bmatrix} = \begin{bmatrix} r_Q \\ r_F \\ r_C \\ r_t \end{bmatrix}. \tag{8.6}$$

(Note that the coefficient matrix is the Jacobian of the nonlinear equations (8.5a).) Different primal-dual methods are obtained from different choices of the right-hand side vector (r_Q, r_F, r_C, r_t). The duality gap μ defined by

$$\mu = \lambda^T t / m \tag{8.7}$$

is typically used as a measure of optimality of the current point (w, π, λ, t). In principle, primal-dual interior-point methods ensure that the norm of the function $\mathcal{F}$ defined by (8.5a) remains bounded by a constant multiple of μ at each iterate, thus ensuring that μ is also a measure of infeasibility of the current point. However, the latter condition is rarely checked in practical algorithms.

We use a variant of Mehrotra's predictor-corrector algorithm (Mehrotra 1992) to solve (8.4). This algorithm has proved to be the most effective approach for general linear programs and is similarly effective for convex quadratic programming. The first part of the Mehrotra search direction—the *predictor* or *affine-scaling* step—is simply a pure Newton step for the system (8.5a), obtained by solving (8.6) with the following right-hand side:

$$\begin{bmatrix} r_Q \\ r_F \\ r_C \\ r_t \end{bmatrix} = -\mathcal{F}(w, \pi, \lambda, t) = -\begin{bmatrix} Qw + F^T\pi + C^T\lambda + c \\ -Fw + g \\ -Cw - t + d \\ T\Lambda e \end{bmatrix}. \tag{8.8}$$

We denote the corresponding solution of (8.6) by $(\Delta w_{\text{aff}}, \Delta \pi_{\text{aff}}, \Delta \lambda_{\text{aff}}, \Delta t_{\text{aff}})$. The second part of the search direction—the *centering-corrector* direction $(\Delta w_{\text{cc}}, \Delta \pi_{\text{cc}}, \Delta \lambda_{\text{cc}}, \Delta t_{\text{cc}})$—is calculated by choosing the centering parameter $\sigma \in [0, 1)$ as outlined below and solving the system (8.6) with the following right-hand side:

$$\begin{bmatrix} r_Q \\ r_F \\ r_C \\ r_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\Delta T_{\text{aff}} \Delta \Lambda_{\text{aff}} e + \sigma \mu e \end{bmatrix}, \tag{8.9}$$

where ΔT_{aff} and $\Delta \Lambda_{\text{aff}}$ are the diagonal matrices constructed from the elements of Δt_{aff} and $\Delta \lambda_{\text{aff}}$, respectively.

The following heuristic for choosing the value of σ has proved to be highly effective. We first compute the maximum step length α_{aff} that can be taken along the affine-scaling direction, as follows:

$$\alpha_{\text{aff}} = \arg\max \{\alpha \in [0,1] \mid (\lambda, t) + \alpha(\Delta\lambda_{\text{aff}}, \Delta t_{\text{aff}}) \geq 0\}.$$

The duality gap μ_{aff} attained from this full step to the boundary is

$$\mu_{\text{aff}} = (\lambda + \alpha \Delta \lambda_{\text{aff}})^T (t + \alpha \Delta t_{\text{aff}})/m.$$

Finally, we set

$$\sigma = \left(\frac{\mu_{\text{aff}}}{\mu} \right)^3 .$$

The search direction is obtained by adding the predictor and centering-corrector directions, as follows:

$$(\Delta w, \Delta \pi, \Delta \lambda, \Delta t) = (\Delta w_{\text{aff}}, \Delta \pi_{\text{aff}}, \Delta \lambda_{\text{aff}}, \Delta t_{\text{aff}}) + (\Delta w_{\text{cc}}, \Delta \pi_{\text{cc}}, \Delta \lambda_{\text{cc}}, \Delta t_{\text{cc}}). \tag{8.10}$$

Note that the coefficient matrix in (8.6) is the same for both the predictor and centering-corrector systems, so just one factorization of this matrix is required at each iteration. Apart from this factorization, the main computational operations at each iteration include two back-substitutions for two different right-hand sides, and a number of matrix-vector operations.

The distance we move along the direction (8.10) is defined in terms of the maximum step $\alpha_{\max}$ that can be taken without violating the condition (8.5b):

$$\alpha_{\max} = \arg\max \{\alpha \in [0,1] \mid (\lambda, t) + \alpha(\Delta \lambda, \Delta t) \geq 0\}.$$

The actual steplength α is chosen to be

$$\alpha \leftarrow \gamma \alpha_{\max}, \tag{8.11}$$

where γ is a parameter in the range $(0, 1)$ chosen to ensure that the pairwise products $\lambda_i t_i$ do not become too unbalanced. The value of γ is typically close to 1; it has proved effective in practice to allow it to approach 1 as the algorithms gets closer and closer to the solution. See Mehrotra (1992) for the details of a heuristic for choosing γ.

The algorithm does not require the initial point to be feasible, and checks can be added to detect problems for which no feasible points exist. In our case, feasibility of the MPC problem obtained from the can be determined a priori by solving a linear program.

Finally, we note that block elimination can be applied to the system (8.6) to obtain reduced systems with more convenient structures. By eliminating Δt, we obtain the following system:

$$\begin{bmatrix} Q & F^T & C^T \\ -F & & \\ -C & & \Lambda^{-1} T \end{bmatrix} \begin{bmatrix} \Delta w \\ \Delta \pi \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} r_Q \\ r_F \\ r_C + \Lambda^{-1} r_t \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} \hat{r}_Q \\ \hat{r}_F \\ \hat{r}_C \end{bmatrix}. \tag{8.12}$$

Since $\Lambda^{-1}T$ is a positive diagonal matrix, we can easily eliminate $\Delta\lambda$ as well to obtain

$$\begin{bmatrix} Q + C^T\Lambda T^{-1}C & F^T \\ -F & 0 \end{bmatrix} \begin{bmatrix} \Delta w \\ \Delta\pi \end{bmatrix} = \begin{bmatrix} r_Q - C^T T^{-1}(\Lambda r_C + r_t) \\ r_F \end{bmatrix}. \tag{8.13}$$

As we see in the next section, these eliminations can be applied to our particular problem to put the system in a form in which we can apply the Riccati block-elimination technique of Sections 8.2.3 and 8.2.4.

We conclude with a note on the sizes of elements in t and λ and their effect on elements of the matrices in (8.12) and (8.13). In path-following interior-point methods that adhere rigorously to the theory, iterates are confined to a region in which the pairwise products $t_i\lambda_i$ are not too different from each other in size. A bound of the form

$$t_i\lambda_i \geq \gamma\mu \tag{8.14}$$

is usually enforced, where μ is the average value of $t_i\lambda_i$ (see (8.7)) and $\gamma \in (0, 1)$ is constant, typically $\gamma = 10^{-4}$. When the primal-dual solution set for (8.4) is bounded, we have further that

$$t_i \leq \beta, \;\; \lambda_i \leq \beta, \;\; i = 1, 2, \ldots, m, \tag{8.15}$$

for some constant bound $\beta > 0$. It follows immediately from (8.14) and (8.15) that

$$\frac{\gamma}{\beta^2}\mu \leq \frac{t_i}{\lambda_i} \leq \frac{\beta^2}{\gamma}\mu^{-1}, \;\; \frac{\gamma}{\beta^2}\mu \leq \frac{\lambda_i}{t_i} \leq \frac{\beta^2}{\gamma}\mu^{-1}. \tag{8.16}$$

Hence, the diagonal elements of the matrices $T^{-1}\Lambda$ and $\Lambda^{-1}T$ lie in the range $[\Theta(\mu), \Theta(\mu^{-1})]$.

Although bounds of the form (8.14) are not enforced explicitly in most implementations of Mehrotra's algorithm, computational experience shows that they are almost always satisfied in practice. Hence, it is reasonable to assume, as we do in the analysis of numerical stability below, that the estimates (8.16) are satisfied by iterates of our algorithm.

8.2.2 Efficient MPC Formulation

The optimal control problem (8.1), (8.2) traditionally has been viewed as a problem in which just the inputs are variables, while the states are eliminated by direct substitution using the transition equation (8.2b) (see, for example, (Muske and Rawlings 1993)). We refer to this formulation hereafter as the standard method. Unfortunately, the constraint and Hessian matrices in the reduced problem resulting from this procedure are generally dense, so the computational cost of solving the problem is proportional to N^3. Efficient commercial solvers for dense quadratic programs (such as QPSOL (Gill, Murray, Saunders and Wright 1983)) can then be applied to the reduced problem.

Unless the number of stages N is small, the $O(N^3)$ cost of the standard method is unacceptable because the "unconstrained" version of (8.1) is known to be solvable in $O(N)$ time by using a Riccati equation or dynamic programming. We are led to ask whether there is an algorithm for the constrained problem (8.1), (8.2) that preserves the $O(N)$ behavior. In fact, the interior-point algorithm of the preceding section almost attains this goal, since it can be applied to the problem (8.1), (8.2) at a cost of $O(N)$ operations *per iteration*. The rows and columns of the reduced linear systems (8.12) and (8.13) can be rearranged to make these matrices *banded*, with dimension proportional to N and and bandwidth independent of N. Since the number of iterations required by the interior-point algorithm depends only weakly on N in practice, the total computational cost of this approach is only slightly higher than $O(N)$. In both the active set and interior-point approaches, the dependence of solution time on other parameters, such as the number of inputs, the number of states, and the number of side constraints, is cubic.

Wright (1993, 1997*a*) describes a scheme in which these banded matrices are explicitly formed and factored with a general banded factorization routine. In the next section, we show that the linear system to be solved at each interior-point iteration can be reduced to a form identical to the "unconstrained" version of (8.1), (8.2), that is, a form in which the side constraints (8.2c), (8.2d) are absent. Hence, a Riccati recursion similar to the technique used for the unconstrained problem can be used to solve this linear system. Even though such a scheme places restrictions on the use of pivoting for numerical stability, we show by a simple argument that numerical stability can be expected.

Suppose that the interior-point algorithm of Section 8.2.1 is applied to the problem (8.1), (8.2). We use λ_k, ζ_k, and η_k to denote the Lagrange multipliers for the constraints (8.2c), (8.2d), and (8.2e), respectively. We rearrange the linear system (8.12) to be solved at each iteration of the interior-point method by "interleaving" the variables and equations according to stage index. That is, the primal and dual variables for stage 0 are listed before those for stage 1, and so on. For this ordering, the rows of the system (8.12) that correspond to stage k are as follows:

$$\left[\begin{array}{cccccccccc}
\cdots & Q & M & -G^T & A^T & & & & & \\
 & M^T & R & D^T & B^T & & & & & \\
 & -G & D & -\Sigma_k^D & & & & & & \\
 & A & B & & & & & & -I & \\
 & & & & & -\Sigma_{k+1}^{\epsilon} & & -I & & \\
 & & & & & & -\Sigma_{k+1}^{H} & -I & H & \\
 & & & & & -I & -I & Z & & \\
 & & & & -I & & H^T & & Q & \cdots
\end{array}\right]
\left[\begin{array}{c}
\vdots \\ \Delta x_k \\ \Delta u_k \\ \Delta\lambda_k \\ \Delta p_{k+1} \\ \Delta\xi_{k+1} \\ \Delta\eta_{k+1} \\ \Delta\epsilon_{k+1} \\ \Delta x_{k+1} \\ \vdots
\end{array}\right]
=
\left[\begin{array}{c}
\vdots \\ r_k^x \\ r_k^u \\ r_k^\lambda \\ r_{k+1}^p \\ r_{k+1}^\xi \\ r_{k+1}^\eta \\ r_{k+1}^\epsilon \\ r_{k+1}^x \\ \vdots
\end{array}\right]. \tag{8.17}$$

In this system, the diagonal matrices Σ_k^D, Σ_k^ϵ, and Σ_k^H, which correspond to $\Lambda^{-1}T$ in the general system (8.12), are defined by

$$\Sigma_k^D = (\Lambda_k)^{-1}T_k^\lambda, \;\; \Sigma_k^\epsilon = (\Xi_k)^{-1}T_k^\xi, \;\; \Sigma_k^H = (\mathcal{H}_k)^{-1}T_k^\eta, \tag{8.18}$$

where Λ_k, Ξ_k, and $\mathcal{H}_k$ are the diagonal matrices whose diagonal elements are the Lagrange multipliers λ_k, ξ_k, and η_k, while T_k^λ, T_k^ξ, and T_k^η are likewise diagonal matrices constructed from the slack variables associated with the constraints (8.2c), (8.2d), and (8.2e), respectively. The final rows in this linear system are

$$\left[\begin{array}{cccccccccc}
\cdots & Q & M & -G^T & A^T & & & & & \\
 & M^T & R & D^T & B^T & & & & & \\
 & -G & D & -\Sigma_{N-1}^D & & & & & & \\
 & A & B & & & & & & -I & \\
 & & & & & -\Sigma_N^{\epsilon} & & -I & & \\
 & & & & & & -\Sigma_N^{H} & -I & H & \\
 & & & & & -I & -I & Z & & \\
 & & & & -I & & H^T & & \bar{Q}_N & F^T \\
 & & & & & & & & F &
\end{array}\right]
\left[\begin{array}{c}
\vdots \\ \Delta x_{N-1} \\ \Delta u_{N-1} \\ \Delta\lambda_{N-1} \\ \Delta p_N \\ \Delta\xi_N \\ \Delta\eta_N \\ \Delta\epsilon_N \\ \Delta x_N \\ \Delta\beta
\end{array}\right]
=
\left[\begin{array}{c}
r_{N-1}^x \\ r_{N-1}^u \\ r_{N-1}^\lambda \\ r_{N-1}^p \\ r_{N-1}^\xi \\ r_{N-1}^\eta \\ r_{N-1}^\epsilon \\ r_N^x \\ r^\beta
\end{array}\right], \tag{8.19}$$

where β denotes the Lagrange multiplier for the endpoint constraint (8.2f).

By eliminating the Lagrange multiplier steps $\Delta\lambda_k$, $\Delta\xi_k$, $\Delta\eta_k$, and $\Delta\epsilon_k$ from the systems (8.17) and (8.19), we derive the following analog of the compact system (8.13):

$$\begin{bmatrix} R_0 & B^T & & & & & & & & & \\ B & & -I & & & & & & & & \\ & -I & Q_1 & M_1 & A^T & & & & & & \\ & & M_1^T & R_1 & B^T & & & & & & \\ & & A & B & & -I & & & & & \\ & & & & -I & Q_2 & M_2 & A^T & & & \\ & & & & & M_2^T & R_2 & B^T & & & \\ & & & & & A & B & \ddots & \ddots & & \\ & & & & & & & \ddots & Q_N & F^T \\ & & & & & & & & F & \end{bmatrix} \begin{bmatrix} \Delta u_0 \\ \Delta p_0 \\ \Delta x_1 \\ \Delta u_1 \\ \Delta p_1 \\ \Delta x_2 \\ \Delta u_2 \\ \vdots \\ \Delta x_N \\ \Delta \beta \end{bmatrix} = \begin{bmatrix} \tilde{r}_0^u \\ \tilde{r}_0^p \\ \tilde{r}_1^x \\ \tilde{r}_1^u \\ \tilde{r}_1^p \\ \tilde{r}_2^x \\ \tilde{r}_2^u \\ \vdots \\ \tilde{r}_N^x \\ r^\beta \end{bmatrix}, \tag{8.20}$$

where

$$\begin{aligned} R_k &= R + D^T(\Sigma_k^D)^{-1}D, & k &= 0,\dots,N-1, \\ M_k &= M - G^T(\Sigma_k^D)^{-1}D, & k &= 1,\dots,N-1, \\ Z_k &= Z + (\Sigma_k^\epsilon)^{-1} + (\Sigma_k^H)^{-1}, & k &= 1,\dots,N, \\ Q_k &= Q + G^T(\Sigma_k^D)^{-1}G + H^T[(\Sigma_k^H)^{-1} - (\Sigma_k^H Z_k \Sigma_k^H)^{-1}]H, & k &= 1,\dots,N-1, \\ Q_N &= \bar{Q}_N + H^T[(\Sigma_N^H)^{-1} - (\Sigma_N^H Z_N \Sigma_N^H)^{-1}]H, \end{aligned} \tag{8.21}$$

and

$$\begin{aligned} \tilde{r}_k^u &= r_k^u + D^T(\Sigma^D)^{-1}r_k^\lambda, & k &= 0,\dots,N-1, \\ \tilde{r}_k^p &= r_k^p, & k &= 0,\dots,N-1, \\ \tilde{r}_k^\epsilon &= r^\epsilon - (\Sigma_k^\epsilon)^{-1}r_k^\xi - (\Sigma_k^H)^{-1}r_k^\eta, & k &= 1,\dots,N, \\ \tilde{r}_k^x &= r_k^x + -G^T(\Sigma_k^D)^{-1}r_k^\lambda + H^T(\Sigma_k^H)^{-1}r_k^\eta + H^T(\Sigma_k^H Z_k)^{-1}\tilde{r}_k^\epsilon, & k &= 1,\dots,N-1, \\ \tilde{r}_N^x &= r_N^x + H^T(\Sigma_N^H)^{-1}r_N^\eta + H^T(\Sigma_N^H Z_N)^{-1}\tilde{r}_N^\epsilon. \end{aligned} \tag{8.22}$$

This matrix has the same form as the KKT matrix obtained from the following problem in which the only constraint (apart from the model equation and initial state) is a final point condition:

$$\min_{u,x} \ \Phi(u,x) = \frac{1}{2}u_0^T R_0 u_0 + \sum_{k=1}^{N-1} \frac{1}{2}(x_k^T Q_k x_k + u_k^T R_k u_k + 2x_k^T M_k u_k) + \frac{1}{2}x_N^T Q_N x_N, \tag{8.23}$$

subject to

$$x_0 = \hat{x}_j, \qquad \text{(fixed)}, \tag{8.24a}$$

$$x_{k+1} = Ax_k + Bu_k, \qquad k = 0,1,\dots,N-1, \tag{8.24b}$$

$$Fx_N = 0. \tag{8.24c}$$

The problem (8.23), (8.24) is convex if the matrices R_0, Q_N, and

$$\begin{bmatrix} Q_k & M_k \\ M_k^T & R_k \end{bmatrix}, \qquad k = 1,2,\dots,N-1, \tag{8.25}$$

are all SPSD. The following brief discussion shows that this property holds. First, we show that

$$(\Sigma_k^H)^{-1} - (\Sigma_k^H Z_k \Sigma_k^H)^{-1} \ \text{ is positive diagonal for all } k = 1,2,\dots,N. \tag{8.26}$$

By using the definition of Z_k above, together with the diagonality of Z and Σ_k^ϵ, we have that

$$\begin{aligned} &(\Sigma_k^H)^{-1} - (\Sigma_k^H Z_k \Sigma_k^H)^{-1} \\ &= (\Sigma_k^H)^{-1}[I - (\Sigma_k^H Z_k)^{-1}] = (\Sigma_k^H)^{-1}[I - \left(\Sigma_k^H Z + \Sigma_k^H(\Sigma_k^\epsilon)^{-1} + I\right)^{-1}]. \end{aligned}$$

Since Z, Σ_k^H, and Σ_k^ϵ are all positive diagonal matrices, the final expression above is a product of two positive diagonal matrices, and therefore is itself positive diagonal. Hence, property (8.26) holds. Note from (8.21) that Q_N is an SPSD modification of an SPSD matrix, and therefore is itself SPSD. Note too that from (8.21) again, we have

$$\begin{bmatrix} Q_k & M_k \\ M_k^T & R_k \end{bmatrix}$$
$$= \begin{bmatrix} Q & M \\ M^T & R \end{bmatrix} + \begin{bmatrix} D^T \\ -G^T \end{bmatrix} (\Sigma_k^D)^{-1} \begin{bmatrix} D & -G \end{bmatrix} + \begin{bmatrix} H^T[(\Sigma_k^H)^{-1} - (\Sigma_k^H Z_k \Sigma_k^H)^{-1}]H & 0 \\ 0 & 0 \end{bmatrix},$$

for all $k = 1, 2, \ldots, N-1$. Because of (8.26), we have that the left-hand side of this expression is a sum of SPSD terms, and therefore is itself SPSD. Finally, note from (8.21) that each R_k, $k = 0, 1, \ldots, N-1$ is the sum of a PSD matrix R and an SPSD term $D^T(\Sigma_k^D)^{-1}D$, and is therefore itself SPD. We conclude that the objective function (8.23) is convex.

If we use n to denote the number of components of each state vector x_k and m to denote the number of components of each input vector u_k, we find that the banded coefficient matrix in (8.20) has dimension approximately $N(2n+m)$ and half-bandwidth approximately $2n+m$, so that the computational cost of factoring it by Gaussian elimination would be proportional to $N(m+n)^3$. This estimate is linear in N, unlike the naive dense implementation for which the cost grows like $N^3(m+n)^3$.

8.2.3 Block Elimination: No Endpoint Constraints

We can improve the efficiency of the algorithm by applying a block factorization scheme to (8.20) in place of the elimination scheme for general banded matrices. In this section, we consider the case in which endpoint constraints are not present in the problem (so that the quantities F, $\Delta\beta$, and r^β do not appear in (8.20)). We describe a block elimination scheme and show that it yields a Riccati recursion.

For simplicity, we rewrite the system (8.20) for the case of no endpoint constraints as follows:

$$\begin{bmatrix} R_0 & B^T & & & & & & & & \\ B & & -I & & & & & & & \\ & -I & Q_1 & M_1 & A^T & & & & & \\ & & M_1^T & R_1 & B^T & & & & & \\ & & A & B & & -I & & & & \\ & & & & -I & Q_2 & M_2 & A^T & & \\ & & & & & M_2^T & R_2 & B^T & & \\ & & & & & A & B & \ddots & \ddots & \\ & & & & & & & & \ddots & \bar{Q}_N \end{bmatrix} \begin{bmatrix} \widehat{\Delta u}_0 \\ \widehat{\Delta p}_0 \\ \widehat{\Delta x}_1 \\ \widehat{\Delta u}_1 \\ \widehat{\Delta p}_1 \\ \widehat{\Delta x}_2 \\ \widehat{\Delta u}_2 \\ \vdots \\ \widehat{\Delta x}_N \end{bmatrix} = \begin{bmatrix} \tilde{r}_0^u \\ \tilde{r}_0^p \\ \tilde{r}_1^x \\ \tilde{r}_1^u \\ \tilde{r}_1^p \\ \tilde{r}_2^x \\ \tilde{r}_2^u \\ \vdots \\ \tilde{r}_N^x \end{bmatrix}. \tag{8.27}$$

Our scheme yields a set of matrices $\Pi_k \in \mathbb{R}^{n\times n}$ and vectors $\pi_k \in \mathbb{R}^n$, $k = N, N-1, \ldots, 1$, such that the following relationship holds between the unknown vectors $\widehat{\Delta p}_{k-1}$ and $\widehat{\Delta x}_k$ in (8.27):

$$-\widehat{\Delta p}_{k-1} + \Pi_k \widehat{\Delta x}_k = \pi_k, \qquad k = N, N-1, \ldots, 1. \tag{8.28}$$

We can see immediately from (8.27) that (8.28) is satisfied for $k = N$ if we define

$$\Pi_N = \bar{Q}_N, \qquad \pi_N = \tilde{r}_N^x. \tag{8.29}$$

The remaining quantities Π_k and π_k can be generated recursively. If (8.28) holds for some k, we can combine this equation with three successive block rows from (8.27) to obtain the following subsystem:

$$\begin{bmatrix} -I & Q_{k-1} & M_{k-1} & A^T & \\ & M_{k-1}^T & R_{k-1} & B^T & \\ & A & B & 0 & -I \\ & & & -I & \Pi_k \end{bmatrix} \begin{bmatrix} \widehat{\Delta p}_{k-2} \\ \widehat{\Delta x}_{k-1} \\ \widehat{\Delta u}_{k-1} \\ \widehat{\Delta p}_{k-1} \\ \widehat{\Delta x}_k \end{bmatrix} = \begin{bmatrix} \tilde{r}_{k-1}^x \\ \tilde{r}_{k-1}^u \\ \tilde{r}_{k-1}^p \\ \pi_k \end{bmatrix}. \tag{8.30}$$

Elimination of $\widehat{\Delta p}_{k-1}$ and $\widehat{\Delta x}_k$ yields

$$\begin{bmatrix} -I & Q_{k-1} + A^T \Pi_k A & A^T \Pi_k B + M_{k-1} \\ 0 & B^T \Pi_k A + M_{k-1}^T & R_{k-1} + B^T \Pi_k B \end{bmatrix} \begin{bmatrix} \widehat{\Delta p}_{k-2} \\ \widehat{\Delta x}_{k-1} \\ \widehat{\Delta u}_{k-1} \end{bmatrix} = \begin{bmatrix} \tilde{r}_{k-1}^x + A^T \Pi_k \tilde{r}_{k-1}^p + A^T \pi_k \\ \tilde{r}_{k-1}^u + B^T \Pi_k \tilde{r}_{k-1}^p + B^T \pi_k \end{bmatrix}. \tag{8.31}$$

Finally, elimination of $\widehat{\Delta u}_{k-1}$ yields the equation

$$-\widehat{\Delta p}_{k-2} + \Pi_{k-1} \widehat{\Delta x}_{k-1} = \pi_{k-1}. \tag{8.32}$$

where

$$\begin{aligned} \Pi_{k-1} &= Q_{k-1} + A^T \Pi_k A - \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(8.33a)} \\ &\quad (A^T \Pi_k B + M_{k-1})(R_{k-1} + B^T \Pi_k B)^{-1}(B^T \Pi_k A + M_{k-1}^T), \\ \pi_{k-1} &= \tilde{r}_{k-1}^x + A^T \Pi_k \tilde{r}_{k-1}^p + A^T \pi_k - \qquad\qquad\qquad\qquad\qquad\qquad \text{(8.33b)} \\ &\quad (A^T \Pi_k B + M_{k-1})(R_{k-1} + B^T \Pi_k B)^{-1}(\tilde{r}_{k-1}^u + B^T \Pi_k \tilde{r}_{k-1}^p + B^T \pi_k). \end{aligned}$$

The equation (8.33a) is the famous discrete-time Riccati equation for time-varying weighting matrices.

The solution of (8.27) can now be obtained as follows. We first set Π_N and π_n using (8.29), and then apply (8.33a) to obtain Π_k and π_k for for $k = N-1, N-2, \ldots, 1$. Next, we combine (8.28) for $k = 1$ with the first two rows of (8.27), we obtain

$$\begin{bmatrix} R_0 & B^T & \\ B & & -I \\ & -I & \Pi_1 \end{bmatrix} \begin{bmatrix} \widehat{\Delta u}_0 \\ \widehat{\Delta p}_0 \\ \widehat{\Delta x}_1 \end{bmatrix} = \begin{bmatrix} \tilde{r}_0^u \\ \tilde{r}_0^p \\ \pi_1 \end{bmatrix}, \tag{8.34}$$

and solve this system for $\widehat{\Delta u}_0$, $\widehat{\Delta x}_1$, and $\widehat{\Delta p}_0$. Next, we obtain from (8.31) and (8.30) that

$$\begin{aligned} \widehat{\Delta u}_k &= (R_k + B^T \Pi_{k+1} B)^{-1} [\tilde{r}_k^u + B^T \Pi_{k+1} \tilde{r}_k^p + B^T \pi_{k+1} - (B^T \Pi_{k+1} A + M_k^T) \widehat{\Delta x}_k], \\ \widehat{\Delta x}_{k+1} &= A \widehat{\Delta x}_k + B \widehat{\Delta u}_k, \qquad k = 1, 2, \ldots, N-1. \end{aligned}$$

Finally, the steps $\widehat{\Delta p}_k$ for $k = N-1, N-2, \ldots, 1$ can be recovered from (8.28) The computational cost of the entire process is $O(N(m+n)^3)$.

The question of stability of this approach is an important one. The block elimination/Riccati scheme just described essentially places restrictions on the pivot sequence, that is, the order in which the elements of the matrix in (8.27) are eliminated. (Note however that pivoting for numerical stability can occur "internally," during the factorization of $(R_{k-1} + B^T \Pi_k B)$ in (8.33a) and (8.33b) for $k = N, N-1, \ldots, 2$.) In other circumstances, pivot restrictions are well known to lead to numerical instability, which manifests itself by blowup of the intermediate quantities that arise during the factorization (by which we mean that the intermediate quantities become much larger than the original data of the problem.) However, in the present case, stability can be established by the simple argument of next few paragraphs.

The coefficient matrix in (8.27) becomes increasingly ill-conditioned near the solution. This feature results from wide variation among the elements of the diagonal matrices Σ_k^D, Σ_k^ϵ, and Σ_k^H defined by (8.18) which, as we see from (8.16), can vary between $\Theta(\mu)$ and $\Theta(\mu^{-1})$, where the duality measure μ approaches zero as the iterates approach the solution. It follows from (8.21) that Q_k, $k = 1, 2, \ldots, N$ has its eigenvalues in the range $[0, \Theta(\mu^{-1})]$, while positive definiteness of R ensures that the eigenvalues of R_k, $k = 0, 1, \ldots, N-1$ lie in an interval $[\Theta(1), \Theta(\mu^{-1})]$. Since we showed earlier that the matrices

$$\begin{bmatrix} Q_k & M_k \\ M_k^T & R_k \end{bmatrix}, \qquad k = 1, 2, \ldots, N-1,$$

are SPSD, we deduce from the comments just made that their eigenvalues too must lie in the range $[0, \Theta(\mu^{-1})]$.

We now show that blowup does not occur during computation of the Riccati matrices Π_k and that, in all cases, their eigenvalues lie in the range $[0, \Theta(\mu^{-1})]$. This is certainly true of the starting matrix Π_N defined by (8.29). For the remaining matrices defined by (8.33a), we assume that our assertion is true for Π_k for some k, and prove that it continues to hold for Π_{k-1}. Note that the matrix

$$\begin{bmatrix} Q_{k-1} & M_{k-1} \\ M_{k-1}^T & R_{k-1} \end{bmatrix} + \begin{bmatrix} A^T \\ B^T \end{bmatrix} \Pi_k \begin{bmatrix} A & B \end{bmatrix}, \tag{8.35}$$

has both terms SPSD, with eigenvalues in the range $[0, \Theta(\mu^{-1})]$. Since Π_{k-1} is the Schur complement of $(R_{k-1} + B^T \Pi_k B)$ in the matrix (8.35), it must be positive semidefinite. (Note that Π_{k-1} is well defined by the formula (8.33a), since $R_{k-1} + B^T \Pi_k B$ is an SPSD modification of the SPD matrix R, and so its inverse is well defined.) Moreover, we can see from (8.33a) that Π_k is obtained by subtracting an SPSD matrix from he SPSD matrix $Q_{k-1} + A^T \Pi_k A$, and so its eigenvalues are bounded above by the eigenvalues of the latter matrix. By combining these observations, we conclude that the eigenvalues of Π_{k-1} lie in the range $[0, \Theta(\mu^{-1})]$, as claimed.

For the vectors π_k, $k = N, N-1, \ldots, 1$, we have from the invertibility of $R_{k-1} + B^T \Pi_k B$ that they are well defined. Moreover, since the smallest eigenvalue of $R_{k-1} + B^T \Pi_k B$ has size $\Theta(1)$, we have from the formula (8.33b) and the estimate $\|\Pi_k\| = O(\mu^{-1})$ from the previous paragraph that $\|\pi_k\| = O(\mu^{-2})$, and so this vector does not blow up with k either. (In fact, a more refined analysis can be used to deduce that $\|\pi_k\| = O(\mu^{-1})$, but we omit the details of this argument here.)

We conclude that numerical instability is not a problem in applying the block elimination/Riccati scheme and that, in fact, we can expect this scheme to be as stable as any general scheme based on Gaussian elimination with pivoting.

It might be expected that the inherent ill conditioning of the system (8.27) may lead to an inaccurate computed solution, even when our numerical scheme is stable. It has long been observed by interior-point practitioners, however, that the computed steps are surprisingly effective steps for the algorithm, even on later iterations on which μ is tiny. This observation has recently found some theoretical support (see Wright (1996, 1997*c*)) but the issues involved are beyond the scope of this chapter.

8.2.4 Block Elimination: Endpoint Constraints

When endpoint constraints are present in the problem, they can be accounted for by adding extra recursions to the scheme of the previous section. We describe this approach below, but first mention an alternative way to handle the problem. The presence of endpoint constraints in the model is often symptomatic of the transition matrix A having eigenvalues outside the unit circle. In these circumstances, it is known that Riccati-based techniques can encounter stability difficulties. These difficulties are ameliorated by the technique of parameterizing the input as $u_k = Lx_k + r_k$, where L is a linear stabilizing feedback gain for (A, B), as mentioned in Section 6.3. Alternatively, we can simply discard the Riccati strategy and instead apply a standard banded Gaussian-elimination scheme with partial pivoting to the system (8.20). Though this approach does not exploit the structure of the problem quite as well as the Riccati strategy, its stability is guaranteed. It can be used as a backup approach if stability problems are encountered with the modified Riccati approach that we now describe.

In the language of linear algebra, our modification of the block-elimination approach proceeds by partitioning the coefficient matrix in (8.20) as

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{12}^T & T_{22} \end{bmatrix},$$

where

$$T_{11} = \begin{bmatrix} R_0 & B^T & & & & & \\ B & & -I & & & & \\ & -I & Q_1 & M_1 & A^T & & \\ & & M_1^T & R_1 & B^T & & \\ & & A & B & \ddots & \ddots & \\ & & & & \ddots & \bar{Q}_N \end{bmatrix}, \quad T_{12} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ F^T \end{bmatrix}, \quad T_{22} = 0. \tag{8.36}$$

We partition the right-hand side and solution of (8.20) correspondingly and rewrite the system as

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{12}^T & T_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix},$$

where $r_2 = r^\beta$ and $y_2 = \Delta\beta$. By simple manipulation, assuming that T_{11} is nonsingular, we obtain

$$[T_{22} - T_{12}^T T_{11}^{-1} T_{12}] y_2 = r_2 - T_{12}^T T_{11}^{-1} r_1, \tag{8.37a}$$
$$y_1 = T_{11}^{-1} r_1 - T_{11}^{-1} T_{12} y_2. \tag{8.37b}$$

We calculate the vector $T_{11}^{-1} r_1$ by using the approach of Section 8.2.3. The other major operation is to find $T_{11}^{-1} T_{12}$, which we achieve by solving the following system:

$$\begin{bmatrix} R_0 & B^T & & & & & & & & \\ B & & -I & & & & & & & \\ & -I & Q_1 & M_1 & A^T & & & & & \\ & & M_1^T & R_1 & B^T & & & & & \\ & & A & B & & -I & & & & \\ & & & & -I & Q_2 & M_2 & A^T & & \\ & & & & & M_2^T & R_2 & B^T & & \\ & & & & & A & B & \ddots & \ddots & \\ & & & & & & & \ddots & \bar{Q}_N \end{bmatrix} \begin{bmatrix} \Phi_0^u \\ \Phi_0^p \\ \Phi_1^x \\ \Phi_1^u \\ \Phi_1^p \\ \Phi_2^x \\ \vdots \\ \Phi_N^x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ F^T \end{bmatrix}. \tag{8.38}$$

The structure of this system is identical to (8.27) except that the right-hand side is now a matrix instead of a vector. As in the preceding section, we seek $n \times n_f$ matrices Ψ_k, $k = N, N-1, \ldots, 1$ (where n_f is the number of rows in F) such that the following relationship holds between Φ_{k-1}^p and Φ_k^x satisfying (8.38):

$$-\Phi_{k-1}^p + \Pi_k \Phi_k^x = \Psi_k, \qquad k = N, N-1, \ldots, 1. \tag{8.39}$$

(Note that Π_k in (8.39) are identical to the matrices generated by the formulae (8.29), (8.33a) of the previous section. This is hardly surprising, since these matrices depend only on the coefficient matrix and not on the right-hand side.) An argument like that of the previous section yields the following recursion for Ψ_k:

$$\begin{aligned}\Psi_N &= F^T, \\ \Psi_{k-1} &= A^T\Psi_k - (A^T\Pi_k B + M_{k-1})(R_{k-1} + B^T\Pi_k B)^{-1}B^T\Psi_k, \quad k = N, N-1, \ldots, 2.\end{aligned}$$

We solve (8.38) by using a similar technique to the one used for (8.27).

We now recover the solution of (8.20) via (8.37). By substituting from (8.36) and (8.38), we find that

$$\begin{aligned}T_{22} - T_{12}^T T_{11}^{-1} T_{12} &= -(\Phi_N^x)^T F^T, \\ r_2 - T_{12}^T T_{11}^{-1} r_1 &= r^\beta - F\widehat{\Delta x}_N,\end{aligned}$$

so that $y_2 = \Delta\beta$ can be found directly by substituting into (8.37a). We recover the remainder of the solution vector from (8.37b) by noting that

$$T_{11}^{-1}r_1 - T_{11}^{-1}T_{12}y_2 = \begin{bmatrix} \widehat{\Delta u}_0 \\ \widehat{\Delta p}_0 \\ \widehat{\Delta x}_1 \\ \widehat{\Delta u}_1 \\ \widehat{\Delta p}_1 \\ \widehat{\Delta x}_2 \\ \vdots \\ \widehat{\Delta x}_N \end{bmatrix} - \begin{bmatrix} \Phi_0^u \\ \Phi_0^p \\ \Phi_1^x \\ \Phi_1^u \\ \Phi_1^p \\ \Phi_2^x \\ \vdots \\ \Phi_N^x \end{bmatrix} \Delta\beta.$$

In the implementation, the recurrences for computing Π_k, Ψ_k, and π_k take place simultaneously, as do the recurrences needed for solving the systems (8.27) and (8.38). The additional cost associated with the n_f endpoint constraints is $O(N(m+n)^2 n_f)$. When $n_f < n$—which is a necessary condition for (8.20) to have a unique solution—the cost of solving the full system (8.20) is less than double the cost of solving the subsystem (8.27) alone by the method of the preceding section.

8.2.5 Hot Starting

Model predictive control solves a sequence of similar optimal control problems in succession. If the model is accurate and disturbances are modest, the solution of one optimal control problem can be shifted one time step forward to yield a good approximation to the solution of the next problem in the sequence. Unfortunately, an approximate solution of this type is not a suitable starting guess for the interior-point method, since it usually lies at the boundary of the feasible region, whereas interior-point methods prefer starting point that *strictly* satisfy the inequalities in the constraint set. Starting points close to the so-called central path are more suitable. In the notation of Section 8.2.1, the characteristics of such points are that their pairwise products $\lambda_i t_i$ are similar in value for $i = 1, 2, \ldots, m$ and that the ratio of the KKT violations in (8.5a)—measured by $\mathcal{F}(z, \pi, \lambda, t)$—to the duality gap μ is not too large. We can attempt to find near-central points by bumping components of the "shifted" starting point off their bound. (In the notation of Section 8.2.1, we turn the zero value of either t_i or λ_i into a small positive value.) A second technique is to use a shifted version of one of the earlier interior-point iterates from the previous problem. Since the interior-point algorithm tends to follow the central path, and since the central path is sensitive to data perturbations only near the solution, this strategy generally produces an iterate that is close to the central path for the new optimal control subproblem.

In the presence of new disturbances, the previous solution has little relevance to the new optimal control problem. A starting point can be constructed from the unconstrained solution, or we can perform a cold start from a well-centered point, as is done to good effect in linear programming codes (see Wright (1997*b*, Chapter 10)).

8.3 Computational Results

To gauge the effectiveness of the structured interior-point approach, we tested it against the "standard" quadratic programming approach, in which the states x_k are eliminated from the problem (8.1), (8.2) by using the model equation (8.2b). A reduced problem with unknowns u_k, $k = 0, 1, \ldots, N-1$ and ϵ_k, $k = 1, 2, \ldots, n$ is obtained. The reduction in dimension is accompanied by filling in of the constraint matrices and the Hessian of the objective. The resulting problem is solved with the widely used code QPSOL (Gill et al. 1983), which implements an active set method using dense linear algebra calculations.

We compared these two approaches on three common applications of the model predictive control methodology.

Example 1: Copolymerization Reactor. Congalidis, Richards and Ray (1986) presented the following normalized model for the copolymerization of methyl methacrylate (MMA) and vinyl acetate (VA) in a continuous stirred tank reactor:

$$G(s) = \begin{bmatrix} \frac{0.34}{0.85s+1} & \frac{0.21}{0.42s+1} & \frac{0.50(0.50s+1)}{12s^2+0.4s+1} & 0 & \frac{6.46(0.9s+1)}{0.07s^2 0.3s+1} \\ \frac{-0.41}{2.41s+1} & \frac{0.66}{1.51s+1} & \frac{-0.3}{1.45s+1} & 0 & \frac{-3.72}{0.8s+1} \\ \frac{0.30}{2.54s+1} & \frac{0.49}{1.54s+1} & \frac{-0.71}{1.35s+1} & \frac{-0.20}{2.71s+1} & \frac{-4.71}{0.008s^2+0.41s+1} \\ 0 & 0 & 0 & 0 & \frac{1.02}{0.07s^2+0.31s+1} \end{bmatrix}.$$

The normalized inputs into the system are the flows of monomer MMA (u_1), monomer VA (u_2), initiator (u_3), and transfer agent (u_4), and the temperature of the reactor jacket (u_5). The normalized outputs of the systems are the polymer production rate (y_1), mole fraction of MMA in the polymer (y_2), average molecular weight of the polymer (y_3), and reactor temperature (y_4). The model was realized in block observer canonical form (Chen 1984) where the dimension n of state after the realization is 18, and the number m of inputs is 5. The model was discretized with a sample period of 1.

The normalized inputs were constrained to be within 10% of their nominal operating steady–state values The tuning parameters were chosen to be $Q = C^T C$ (where C is the measurement matrix obtained from the state space realization), while $M = 0$, $R = (0.1)I$, and the number of stages N is 100. Due to the very slow dynamics of the reactor, $\bar{Q}$ was obtained using the technique described by Rawlings and Muske (1993). The parameters z and Z are vacuous, since there are no soft constraints on the state. The controller was simulated with the following state disturbance:

$$[x_0]_j = 0.02 * \sin j.$$

The interior-point method required 14 iterations to solve the optimization problem. Figure 8.1 shows the optimal control profile normalized with the upper bounds on the input constraints.

Example 2: Gage Control of a Polymer Film Process. We considered the gage (cross-directional control) of a 26-lane polymer film process with 26 actuators. We used the following model for our simulation:

$$A = 0.9I, \qquad B = (I - A) * K,$$

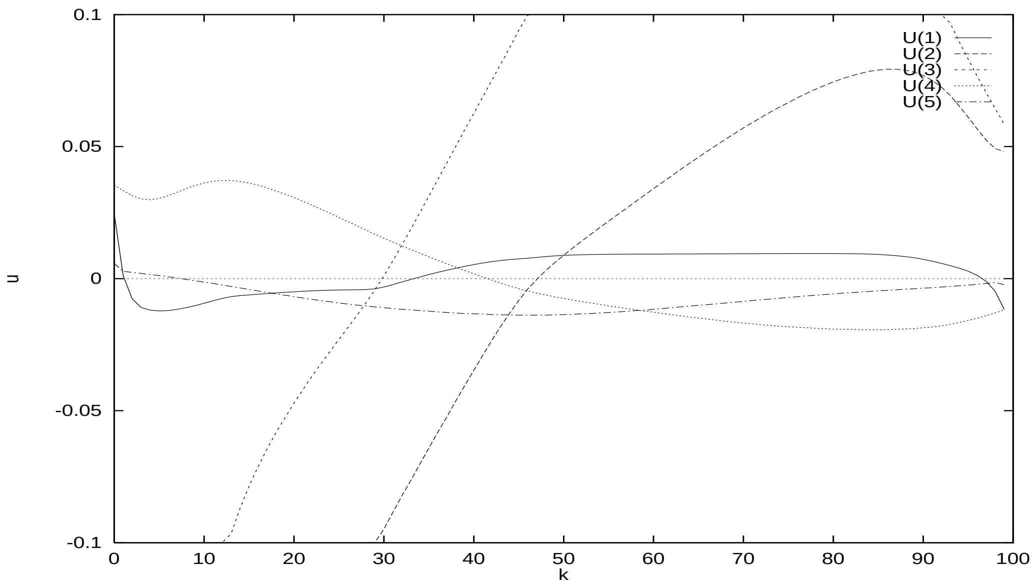


Figure 8.1: Input Profile for Example 1

where the steady-state gain matrix K was extrapolated from data obtained from a 3M polymer film pilot plant. For this example, the dimension of the state n is 26, and the number m of inputs is 26. The state $[x]_j$ denotes the deviated film thickness in the jth lane, and the input $[u]_j$ denotes the deviated position of the jth actuator.

The actuators were constrained between the values of 0.1 and -0.1, while the velocity of the actuators was constrained between the values of 0.025 and -0.025. Since a large difference between actuator positions can create excess stress on the die, we imposed the following restriction on the change in input from stage to stage:

$$|[u]_j - [u]_{j-1}| < 0.05, \qquad j = 2, 3, \ldots, m.$$

We chose the tuning parameters to be

$$Q = I, \;\; R = I, \;\; S = I.$$

The matrix $\bar{Q}$ was obtained from the solution of (6.19). The parameters z and Z are vacuous, since there are no soft constraints on the state. We chose a horizon of $N = 30$ to guarantee that the constraints were satisfied on the infinite horizon. The interior-point method required 11 iterations. Figure 8.2 shows the calculated optimal input profiles.

Example 3: Evaporator. Ricker et al. (1988) presented the following model for an evaporation process in a kraft pulp mill:

$$G(s) = \begin{bmatrix} \frac{1}{30s+1} & 0 \\ \frac{648s}{(30s+1)(20s+1)} & \frac{2.7(-6s+1)}{(20s+1)(5s+1)} \\ \frac{-90s}{(30s+1)(30s+1)} & \frac{-0.1375(-4s+1)}{(30s+1)(2.6s+1)} \end{bmatrix}.$$

The normalized outputs of the process are the feed level (y_1), product concentration (y_2), and product level (y_3). The normalized inputs for the the process are the feed level setpoint (u_1) and the steam flow

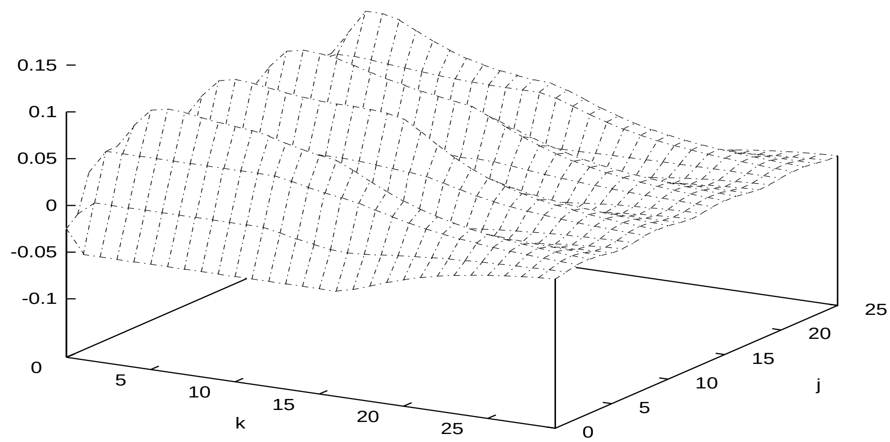


Figure 8.2: Input Profile for Example 2

(u_2). The process was realized in block observer canonical form (Chen 1984) and sampled every 0.5 minutes. The dimension n of the state after the realization is 9, and the number m of input is 3.

Both inputs were constrained to lie in the range $[-0.2, 0.2]$, while the three outputs were constrained to lie in $[-0.05, 0.05]$. A bound of 0.05 was also imposed on the input velocity. The controller was tuned with

$$Q = I, \;\; R = I, \;\; Z = 0, \;\; N = 60.$$

The matrix $\bar{Q}$ was obtained from the solution of (6.19). A constant ℓ_1 penalty of 1000 was sufficient to force the soft constraints to hold when the solution is feasible. We simulated the controller with the following state disturbance:

$$[x_0]_j = \sin(j) + \cos(j).$$

The interior-point method required 18 iterations to solve the optimization problem. Figure 8.3 shows the calculated optimal input profile, while Figure 8.4 shows the predicted output profile. Note that the constraints for y_2 and y_3 are initially violated. The constraint for y_2 is feasible when $k \geq 8$ and the constraint for y_3 is feasible when $k \geq 34$. Increasing the ℓ_1 penalty did not change the resulting solution. Decreasing the ℓ_1 penalty leads to less aggressive control action, but the constraints are violated for a longer duration.

The computational times required by the structured interior-point approach and the naive quadratic programming approach are shown in Table 8.1. Our platform was a DEC Alphastation 250, and the times were obtained with the Unix `time` command. We used the value $\gamma = 0.995$ in (8.11) as the proportion of maximum step to the boundary taken by our algorithm.

For the chosen (large) values of the horizon parameter N, the structured interior-point method easily outperforms the naive quadratic programming approach. For the latter approach, we do not include the time required to eliminate the states. These times were often quite significant, but they are calculated offline. For small values of the horizon parameter N, the naive quadratic programming approach outperforms the structured interior-point method, since the bandwidth is roughly the same relative order of magnitude as the dimensions of (8.20).

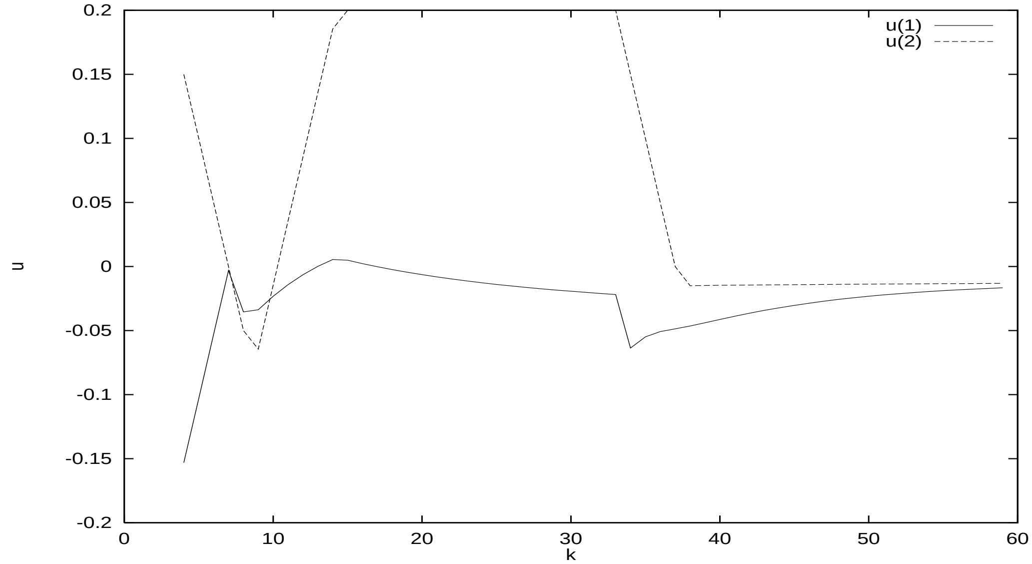


Figure 8.3: Input Profile at $t = 0$ for Example 3

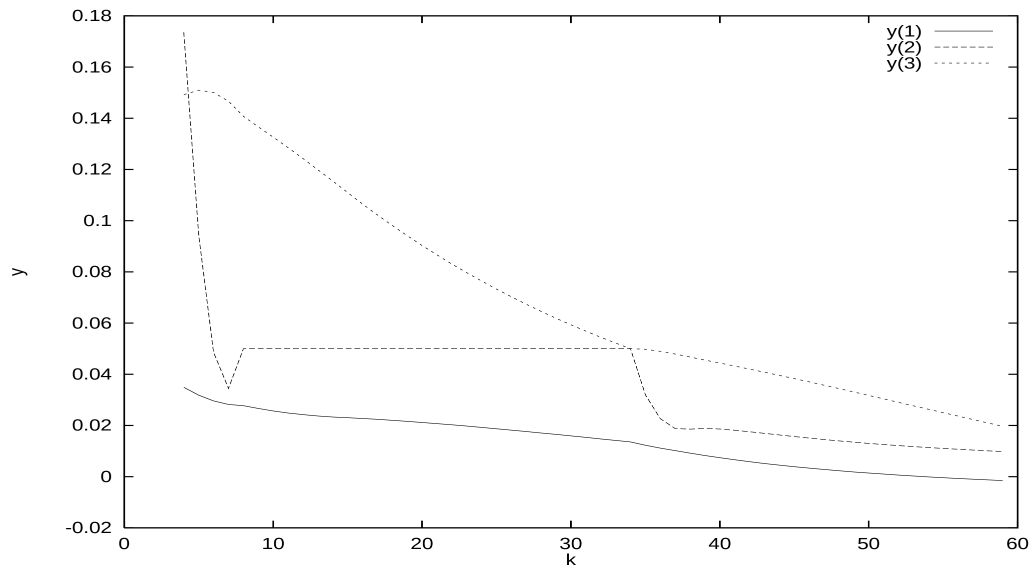


Figure 8.4: Predicted Output Profile for Example 3

Table 8.1: Computational Times (sec)

Example	Structured Interior-Point	Naive Quadratic Programming
1	3.80	23.78
2	20.33	276.91
3	2.01	25.32

8.4 Exploiting Structure in Sheet and Film Forming Processes

Sheet and film forming processes often have large numbers of inputs and outputs. Film forming processes with 30 actuators and more than 150 lane thickness measurements are not uncommon. The processes that produce paper can have 100 actuators and over 400 lane measurements (Wilhelm and Fjeld 1983, Kristinsson and Dumont 1996). Figure 8.5 details the layout of a generic sheet or film forming process. The control objective is to regulate the gage or thickness of the film along the cross–direction. These types of processes produce some challenging computational problems for the control engineer who wishes to use constrained model predictive control to improve product quality. The advantages in using model predictive control lie mainly with the constraint handling. Based on a process model, the optimal control strategy will honor absolute constraints such as actuator travel limits as well as equipment preserving constraints such as spatial travel limits between adjacent actuators to avoid die lip damage.

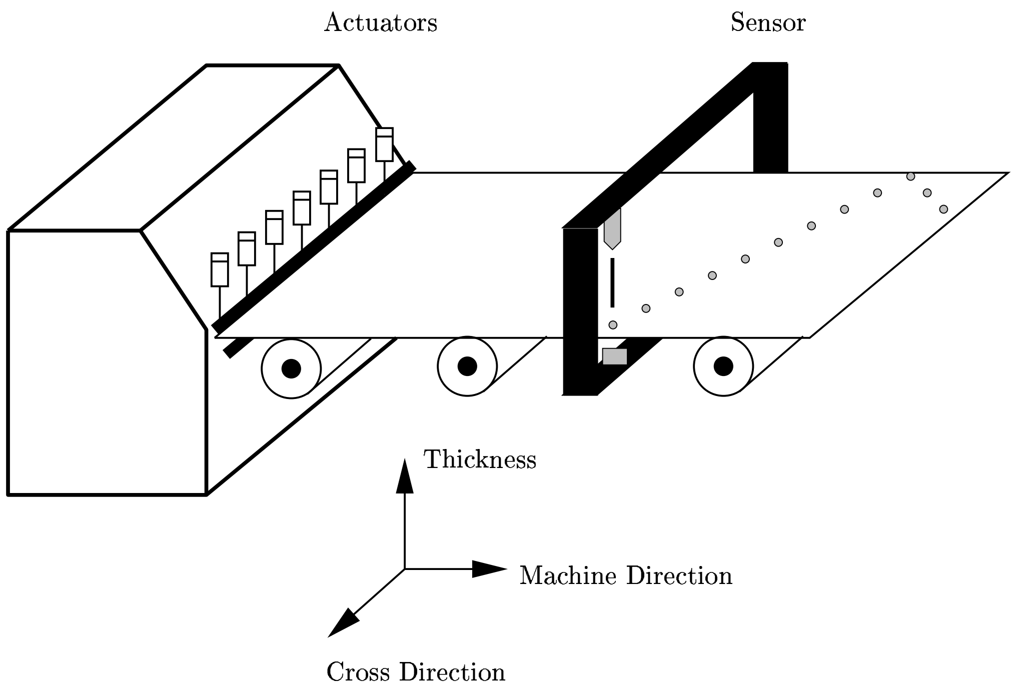


Figure 8.5: Schematic of film or sheet forming process.

The improved efficiency in the structured approach is achieved by tailoring optimization directly to model predictive control. In particular, the structured approach identifies the staged structure of the objective function and constraints, which is common to all model predictive control problems. By further restricting the class of optimization problems, we can expect further increases in computational efficiency. One particular avenue for sparsity exploitation is the structure of the model. For sheet and film forming processes, this structure includes localized spatial dynamics and large time delays.

8.4.1 Exploiting Model Structure

In the structured formulation of the optimization problem discussed in the previous section, the model class is restricted only to finite-dimension, linear time-invariant systems. By restricting our attention to the cross-directional control of sheet and film forming process, we can further restrict the class of models, which allows us to further structure the optimization for improved computational performance. The two distinctive characteristics of the model structure for film and sheet forming processes are localized spatial dynamics and the large input time delay (Braatz, Tyler, Morari, Pranckh and Sartor 1992, Campbell and Rawlings 1996).

We can compactly represent the time delay as the following discrete time state space model

$$x_{k+1} = Ax_k + Bu_{k-d} \tag{8.40}$$

It is straight forward to incorporate time delay in standard linear state space form by expanding the state vector to incorporate the delayed inputs. However, for the structured optimization, it is preferable to work directly with (8.40). If the dynamics of the process are relatively fast with respect to the length of the time delay d, the process dynamics can be approximated as pure delay ($A \equiv 0$) (Braatz et al. 1992, Campbell and Rawlings 1996). The corresponding process model is then

$$x_{k+1} = Pu_{k-d}$$

where P is the steady–state gain of the process.

If there are no dynamics in the process model, then the solution to the optimal control problem would simply be the deadbeat control law $u_k = u_s$ for all k, where u_s is the steady–state input target. The addition of constraints on input and states would not alter the steady-state solution with the exception of input velocity constraints. Letting u_{-1} be the implemented input at time $j-1$ and Δu be the rate of change constraint, the solution to the steady-state control problem is

$$u_k = \min\{\text{sgn}(p_k - u_{-1})p_k, \text{sgn}(p_k - u_{-1})u_s\}$$

where

$$p_k = u_{-1} - \text{sgn}(u_{-1} - u_s)(k+1)\Delta u$$

It is therefore not necessary to use quadratic programming to solve (6.11) when there are no dynamics and $M \equiv 0$ (i.e. no input velocity penalty). The only necessary computation will be the steady–state target calculation (see Section 6.2), the cost of which is minimal.

In the case where process dynamics cannot be ignored or input velocity penalties are desired, it is necessary to use quadratic programming to solve (6.11). We can minimize the computational cost of the quadratic program by exploiting the structure of the state transition matrix and the input distribution matrix. The dominant costs in the solution of Riccati equation (8.33a) are the Cholesky factorization of $B^T\Pi_k B + R_{k-1}$ and the matrix multiplications. There is no simple way to exploit the structure in the Cholesky factorization. However, we can significantly reduce the cost of the matrix multiplications.

The flop count for multiplying two dense matrices $U \in \mathbb{R}^{m\times r}$ and $V \in \mathbb{R}^{r\times n}$ is approximately $2mrn$ (Golub and Van Loan 1983). However, the localized spatial dynamics and also the inherent symmetry of sheet and film forming process (Featherstone and Braatz 1995) gives rise to sparsity and structure which we can exploit in solving (8.33a). Structures that have been observed in the problem data include the following:

1. Diagonal state transition matrices (A)
2. Diagonal penalty matrices (Q, R, M)
3. Diagonal or banded constraints (D, G)
4. Banded, centrosymmetric or Toeplitz gain matrix (P)

By exploiting the sparsity appropriately to avoid multiplications by zero, we can expect the computational cost to be subcubic in the state and control dimensions for a large class of sheet and film forming processes.

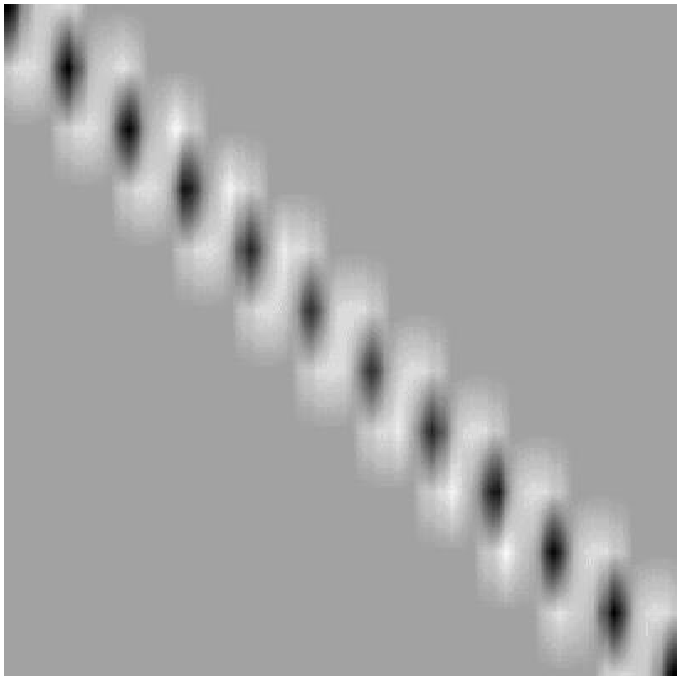


Figure 8.6: Structure of Steady State Gain Matrix P

8.4.2 Computational Results

To gauge the potential benefits of exploiting the structured sparsity of sheet and film forming processes, we consider three computational examples. In the first example, we investigate a process with 100 outputs and 50 inputs with saturation constraints on the input. For the second example, we investigate a process with 100 outputs and 100 inputs with bound constraints on the input. In the third example, we investigate the same process as the second example with the addition of the following spatial constraint

$$|[u]_j - [u]_{j-1}| < b, \;\; j \in [1, m]$$

which is used to reduce excess stress on the die lip caused by large differences between the position of adjacent actuators. In all three examples the following tuning parameters are used: $Q = I$, $R = I$, and $S = 0$. The control objective for all three examples is to reject a sinusoidal state disturbance.

For all three simulations we assume a state space model of the following structure

$$A = 0.95I, \;\; B = (I - A) * P,$$

in which the steady–state gain was extrapolated from a smaller model identified from a polymer film pilot plant. While this model is not representative of all sheet and film forming process, it does possess some of the characteristics features. The structure of the gain matrix for the 100×100 examples is shown as the shaded picture in Figure 8.6. The gain matrix has an average bandwidth of 20. The tight band is expected since it would be unlikely that a single actuator would be capable of influencing the thickness of the entire sheet. Figures 8.7 and 8.8 show the calculated inputs and states for the first example.

Table 8.2 shows the computational times required for a single calculation of quadratic program in (6.11) for the three examples with different control horizon lengths N using the commercial package QPSOL (version 3.2) (Gill et al. 1983) as a representative dense optimization approach, the structured interior point method (SIP) mentioned in Section 2, and a sparse structured interior point method

Example	N	QPSOL	SIP	SSIP
4	10	50.9	39.1	26.7
	20	410.7	75.0	54.6
5	10	340.7	80.3	43.8
	20	–	160.8	89.5
6	10	470.2	110.9	48.4

Table 8.2: Comparison of Computational Times (sec)

(SSIP) which, as its name suggests, exploits the structure and sparsity in the optimization data during solution of the Riccati equation (8.33a). The computational times were obtained using the Unix `time` command on an Alphastation 250. For the QPSOL experiments, the listed times do not account for the time required to construct the optimization data. These times were often significant, but they represent off-line calculations.

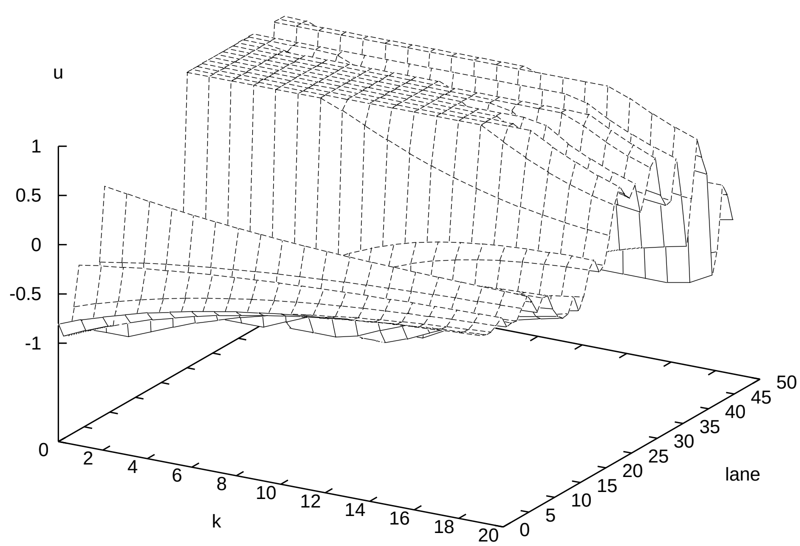


Figure 8.7: Optimal Input Profile for Example 4

As expected there are significant improvements in computational performance obtained by tailoring the optimization to the specific problem. The performance improvement increases with the size of the problem. The computational cost increases linearly as expected for the structured approaches while the cost increases cubically for the dense approach. For smaller problems (less than 100 decision variable), QPSOL tends to perform or outperform the structured methods, since the structure of the problem dominates the optimization algorithm only for problems of large scale.

The sparse structured approach is almost twice as fast as the structured approach for the three examples, while it was almost an order of magnitude faster than the dense approach for the third example. The performance improvements were obtained by recognizing the diagonal structure of A and R, the banded structure of B, and the diagonal/banded structure of the input constraint matrix D.

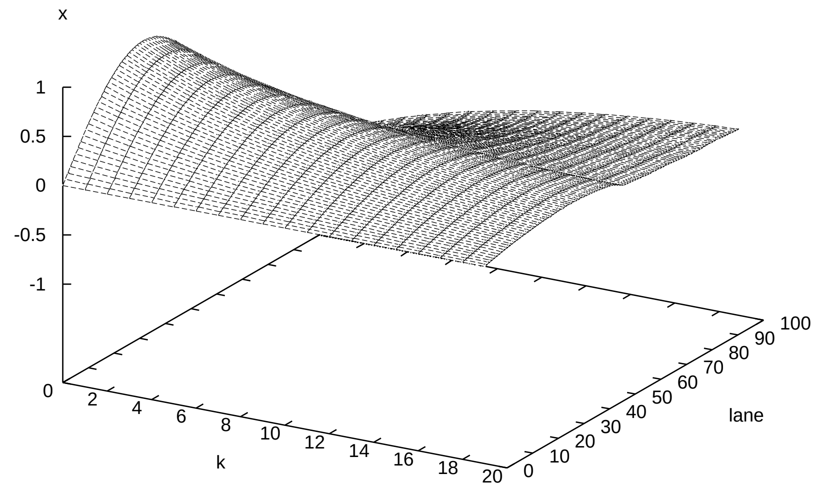


Figure 8.8: Predicted State Profile for Example 4

The changes to the structured optimization code only consisted of augmenting the matrix–matrix and matrix-vector multiplication routines to account for the sparse matrix structures.

For the second example with a control horizon of 20, we were unable to construct the Hessian for QPSOL due to memory requirements. In contrast, the memory requirements for the structured approaches were well within our computation limits. The structured approach required data structures on order of ten thousand elements as compared to data structures on order of one million elements required for the dense approach.

8.5 Concluding Remarks

We conclude with four brief comments on the structured interior-point method for MPC. The first is that the structured method presented is also directly applicable to the dual problem of MPC, the constrained moving horizon estimation problem. In fact, the estimation problem will provide greater justification for structured approach because long horizons N arise frequently in this context. However, we did not investigate applying the structured optimization approach because the theory for linear constrained receding horizon estimators is still in its infancy.

The second comment is that we can extend the structured method to nonlinear MPC by applying the approach of this chapter to the linear-quadratic subproblems generated by sequential quadratic programming. Wright (1993), Arnold et al. (1994), and Steinbach (1994) all apply a similar technique to discrete-time optimal-control problems. While some theory for nonlinear MPC is available, the questions of robust implementation and suitable formulation of nonlinear MPC have not been resolved. See Mayne (1997) for a discussion of the some of the issues.

Third, since the computational cost of the proposed algorithm is $O(N(m+n)^3)$, systems with large numbers of states and inputs can still present formidable computational challenges. Since large systems tend to be sparse (that is, A and B tend to be sparse, while Q and R tend to be nearly diagonal), we expect substantial increases in computational performance by exploiting the sparsity in (8.20) through the use of sparse matrix solvers. Since the sparsity tends to be structured in many applications as we demonstrated in Section 8.4, different strategies are preferable for different classes of processes.

The fourth comment concerns time delays, which occur when more than one sampling period elapses before an input u_k affects the state of the system. In the simplest case, we can rewrite the state equation (6.15a) as

$$x_{k+1} = Ax_k + Bu_{k-d}, \tag{8.41}$$

for the case in which the delay is d sampling periods. The natural infinite horizon LQR objective function for this case is

$$\Phi = \frac{1}{2}\sum_{k=0}^{\infty}(x_k^T Q x_k + u_k^T R u_k + 2x_{k+d}^T M u_k), \tag{8.42}$$

where the cross-penalty term relates u_{k-d} and x_k. Since the first $(d+1)$ state vectors $x_0, x_1, \ldots, x_d$ are independent of the inputs, the decision variables in the optimization problem are $x_{d+1}, x_{d+2}, \ldots$ and $u_0, u_1, \ldots$. By defining

$$\tilde{x}_k = x_{k+d}, \qquad k = 0, 1, 2, \ldots,$$

and removing constant terms from (8.42), the objective function and state equation become

$$\Phi' = \tfrac{1}{2}\textstyle\sum_{k=0}^{\infty}(\tilde{x}_k^T Q \tilde{x}_k + u_k^T R u_k + 2\tilde{x}_k^T M u_k), \tag{8.43}$$

$$\tilde{x}_0 = x_d, \;\; \tilde{x}_{k+1} = A\tilde{x}_k + Bu_k, \;\; k = 0, 1, 2, \ldots. \tag{8.44}$$

These formulae have the same form as (6.14) and (6.15).

If no additional constraints of the form (6.15b) are present, a Riccati equation may be used to solve (8.43), (8.44) directly, as in Section 6.3. If state constraints of the form $Hx_k \le h$ or jump constraints of the form $-\Delta_u \Delta u_k \le \Delta_u$ are present (as in (6.11)), we can still apply constraint softening (Section 6.3.2) and use the approaches described in Section 6.3 To obtain finite-horizon versions of (8.43), (8.44). The techniques of Section 8.2.1 can then be used to solve the problem efficiently.

Difficulties may arise, however, when multiple time delays are present, since these may reduce the locality of the relationships between the decision variables and lead to significant broadening of the bandwidths of the matrices in (8.17) and (8.20). A process in which two time delays are present (of d_1 and d_2 sampling intervals) can be described by a state equation of the following form:

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} u_{k-d_1}^1 \\ u_{k-d_2}^2 \end{bmatrix}.$$

A problem with these dynamics can be solved by augmenting the state vector x_k with the input variables $u_{k-d_1}, u_{k-d_1-1}, \ldots, u_{k-d_2+1}$ (assuming that $d_2 > d_1$) and applying the technique for a single time delay outlined above. Alternatively, the KKT conditions for the original formulation can be used directly as the basis of an interior-point method. The linear system to be solved at each interior-point iteration will contain not only diagonal blocks of the form in (8.17), but also a number of blocks at some distance from the diagonal. Some rearrangement to reduce the overall bandwidth may be possible, but expansion of the bandwidth by an amount proportional to $(d_2 - d_1)m$ is inevitable.

Of course, we can also revert to the original approach of eliminating the states $x_0, x_1, \ldots$ from the problem to obtain a problem in which the inputs $u_0, u_1, \ldots$ alone are decision variables. The cost of this approach, too, is higher than in the no-delay case, because the horizon length N usually must be increased to incorporate the effects of the delayed dynamics. One could postulate that certain processes would be effectively handled by the standard approach while others would be effectively handled by the structured approach. Perhaps the only solution is to exercise engineering judgment to decompose the full control problem into smaller problems without large delays and treat the neglected delays connecting the decomposed systems as disturbances. This issue remains unresolved and is a topic of current research.

Chapter 9

Constrained Linear Disturbance Attenuation

9.1 Introduction

As we have illustrated in the previous chapters, the theory of model predictive control (MPC) is relatively mature; many practical and theoretical issues have been resolved. There also exists a host of survey papers on MPC; examples include (García et al. 1989), (Kwon 1994), (Mayne 1997), (Lee and Cooley 1997), and (Mayne, Rawlings, Rao and Scokaert 1999), as well as the book by Camacho and Bordons (1998). In the process industries, MPC is a popular tool for advanced control and thousands of applications have been reported (c.f Qin and Badgwell (1997, 1998)). While the popularity of MPC is widely of documented (particularly in process control), many important issues still have not been satisfactorily resolved. One of these issues is robustness. Numerous authors have addressed the issue of robustness in MPC. An incomplete list comprises of the following articles: (Campo and Morari 1987), (Allwright and Papavasiliou 1992), (Zheng and Morari 1993), (Genceli and Nikolaou 1993), (Kothare, Balakrishnan and Morari 1996), (Lee and Yu 1997), (Badgwell 1997), (Chen, Scherer and Allgöwer 1997), (Magni 1998), (Chen, Scherer and Allgöwer 1998), and (Scokaert and Mayne 1998). While these articles present promising results, they all suffer from two drawbacks. None of the articles listed above discuss stabilizing strategies for output feedback in a state-space setting, and all require the solution of a minimax problem in real-time, a computationally demanding problem that often is not well-posed. Some exceptions that bypass the need for a minimax formulation and rely on alternative formulations not based on worst-case analysis are (Michalska and Mayne 1993), (Polak and Yang 1993*a*), (De Nicolao, Magni and Scattolini 1996), (Magni and Sepulchre 1997), (Santos and Bieger 1999), and (Zheng 1999).

Of importance in the context of this chapter is the constrained disturbance attenuation problem, or $\mathcal{H}_\infty$ problem (Chen et al. 1997, Chen et al. 1998, Magni 1998)), where MPC is formulated as a dynamic game. Using a moving horizon approximation, MPC avoids the solution of a Hamilton-Jacobi-Bellman-Isaacs equation by repetitively solving an open-loop dynamic game. One strength of the dynamic game approach is the relationship to $\mathcal{H}_\infty$ control (c.f. Başar and Bernhard (1995)). This relationship allows one potentially to draw on results and intuitions from $\mathcal{H}_\infty$ control. Furthermore, many robustness problems can be embedded in the disturbance attenuation problem (i.e. the $\mathcal{H}_\infty$ problem) using the classic $M - \Delta$ structure (e.g. Skogestad and Postlethwaite (1996)).

In this chapter, we examine the disturbance attenuation problem in linear MPC using a dynamic game approach. However, unlike previous work, we address two unanswered issues: the real-time solution of minimax problems and output feedback. By examining sufficient conditions for saddle-point solutions, we reformulate the infinite-horizon dynamic game as a finite-dimension quadratic problem. Using forward/backward dynamic programming, we establish a separation principle, thereby solving the output feedback problem. Our work relies heavily on the results of Başar and Bernhard (1995). Their work motivated and provided many of tools necessary for our investigations into the constrained problem.

The chapter is organized as follows. We begin by first defining the disturbance attenuation problem in Section 9.2. In Section 9.3 we formulate the disturbance attenuation problem as a dynamic game. In Section 9.4 we provide a receding horizon formulation for the disturbance attenuation problem. In Section 9.5 we discuss the perturbed reachability problem and then characterize the admissible class of disturbances in Section 9.6. We conclude by summarizing our results in Section 9.7 and then discussing some limitations of MPC and open-loop feedback control in Section 9.8.

9.2 Problem Statement

Consider the closed-loop response of the following discrete-time linear system

$$x_{k+1} = Ax_k + Bu_k + Gw_k, \tag{9.1a}$$

$$y_k = Cx_k + Ev_k, \tag{9.1b}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{U} \subseteq \mathbb{R}^m$, $w_k \in \mathbb{R}^q$, $y_k \in \mathbb{R}^p$, and the matrix E is non-singular, subject to feedback law

$$\mu_k(y_0, \dots, y_{k-1}, \mu_0, \dots, \mu_{k-1}) = u_k.$$

We assume the pair (A, B) is controllable, the pair (A, C) is observable, and the set $\mathbb{U}$ is closed, convex, and $0 \in \text{int } \mathbb{U}$. State constraints are ignored due to the host of technical difficulties regarding feasibility. Because the inputs are constrained ($\mathbb{U} \neq \mathbb{R}^m$), the controller **cannot** necessarily stabilize the entire state space $\mathbb{R}^n$. Rather, a feedback controller may stabilize only a subset of the state space $\mathcal{X} \subseteq \mathbb{R}^n$ (for stable systems, the entire state space is stabilized trivially). To account for the inherent physical limitations of the controller, we need to restrict the class of admissible disturbances. Given $\gamma > 0$, let $\mathcal{D}(\gamma) \subset \mathbb{R}^n \times l_2(\mathbb{R}^{q+p})$ denote the admissible class of disturbances, where $\mathcal{D}(\gamma_1) \subset \mathcal{D}(\gamma_2)$ when $\gamma_2 < \gamma_1$ and

$$\lim_{\gamma \to 0} \mathcal{D}(\gamma) = \mathbb{R}^n \times l_2(\mathbb{R}^{q+p}).$$

We say the disturbance sequence is **admissible** if $(x_0, \{w_k, v_k\}_{k=0}^{\infty}) \in \mathcal{D}(\gamma)$. A system is said to have **finite l_2-gain** $\leq \gamma$ if there exists a constant γ such that for all $N \geq 0$ and admissible disturbances, we have

$$\sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k \leq \gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + v_k^T v_k + x_0^T \Pi^{-1} x_0 \right),$$

where the matrices Q, R, and Π are symmetric positive definite. We say a system has **minimal l_2-gain** $\leq \gamma$ if the system has finite l_2-gain $\leq \gamma$, and there exists no $\gamma^* < \gamma$ such that the system has finite l_2-gain $\leq \gamma^*$. We define the **constrained disturbance attenuation** problem as determining the feedback law

$$\mu_k(y_0, \dots, y_{k-1}, \mu_0, \dots, \mu_{k-1}) = u_k$$

that achieves minimal finite l_2-gain closed-loop performance and satisfies the constraints: $u_k \in \mathbb{U}$. Note that the controller is nonlinear even though the model (9.1) is linear due to the addition of the inequality constraints. As a result, frequency domain analysis is not applicable to the constrained disturbance attenuation problem. Rather, we need to rely on tools from nonlinear analysis, in particular dissipative systems theory and game theory.

9.2.1 Notation

The set of positive real numbers is denoted by $\mathbb{R}_+$. Let $x(k; z, \{u_j\}, \{w_j\})$ denote the solution of the system (9.1) at time k subject to the initial condition $x_0 = z$ at time 0, input sequence $\{u_j\}_{j=0}^{k-1}$, and disturbance sequence $\{w_j\}_{j=0}^{k-1}$. Let $\|x\| = \sqrt{x^T x}$ and $l_2(\mathbb{R}^n)$ denote the space of all sequences $\{a_k\}$ in $\mathbb{R}^n$ for which $\sum_{k=0}^{\infty} \|a_k\|^2 < \infty$. For $\epsilon > 0$, $B_\epsilon := \{x \ : \ \|x\| \leq \epsilon\}$. Let $a \vee b := \max\{a, b\}$.

9.3 Dissipative Systems and Dynamic Games

The theory of dissipative systems (Willems 1972) provides a general framework for analyzing the disturbance attenuation problem for nonlinear systems. From a design perspective, dissipative systems theory allows one to recast the disturbance attenuation as a dynamic game (e.g. van der Schaft (1996)). Because of our interest in a constructive strategy for output feedback, many of our ideas are motivated by the work of James and Baras (1995). While they use dissipative systems theory to establish a closed-loop system has finite l_2-gain, their approach is grounded in the notion of an information state, and in many ways analogous to the work of Bernhard (1995). In this section we discuss infinite-horizon MPC.

Fix $\gamma > 0$ and consider the following **open-loop** zero-sum infinite-horizon game at time T

$$\mathcal{P}(T): \qquad \phi_T^* = \min_{\{u_k\}_{k=0}^{\infty}} \max_{\left(x_0, \{w_k, v_k\}_{k=0}^{\infty}\right)} \{\phi(\{u_k\}, (x_0, \{w_k, v_k\})) \ : \ u_k \in \mathbb{U}, \ \ (x_0, \{w_k, v_k\}) \in \mathcal{D}(\gamma)\}$$

subject to the history constraint, for $k \leq (T-1)$,

$$u_k = u_{k|k}^*, \tag{9.2a}$$
$$y_k = Cx_k + v_k. \tag{9.2b}$$

where the objective function is given by

$$\phi(\{u_k\}, (x_0, \{w_k, v_k\})) := \frac{1}{2} \sum_{k=0}^{\infty} x_k^T Q x_k + u_k^T R u_k - \gamma^2 \left(w_k^T w_k + v_k^T v_k + x_0^T \Pi^{-1} x_0\right)$$

and $x_k := x(k; x_0, \{u_j\}, \{w_j\})$. The solution to $\mathcal{P}(T)$, assuming it exists, are the sequences $\{u_{k|T}^*\}_{k=0}^{\infty}$ and $\left(x_{0|T}^*, \{w_{k|T}^*, v_{k|T}^*\}_{k=0}^{\infty}\right)$. Feedback is introduced by solving the dynamic game $\mathcal{P}(T)$ at each time T, and implementing only the input $u_{T|T}^*$. When the measurement y_T becomes available, we solve $\mathcal{P}(T+1)$ and then repeat the process. The feedback law is given by $\mu_k(\cdot) = u_{k|k}^*$. The history constraints are important in our formulation, because they allow us to introduce output feedback into the problem. At time T, the controller has observed the system response $\{y_k\}_{k=0}^{T-1}$ and, as a result, the class of admissible disturbances are restricted.

Proceeding in an informal manner, let us begin by assuming for simplicity that $\mathcal{D}(\gamma) = \mathbb{R}^n \times l_2(\mathbb{R}^{q \times p})$ and $\mathbb{U} = \mathbb{R}^n$. Due to the open-loop structure of the game $\mathcal{P}(T)$, a solution may not exist. This limitation holds often regardless of our choice of $\mathcal{D}(\gamma)$. If we assume the infinite-horizon problem is the limit of maximums over finite sequences $\left(x_0, \{w_k, v_k\}_{k=0}^{N-1}\right)$ and rely on standard linear quadratic theory for the existence of a minimum, then

$$\phi_T^* \equiv \lim_{N \to \infty} \min_{\{u_k\}_{k=0}^{\infty}} \max_{\left(x_0, \{w_k, v_k\}_{k=0}^{N-1}\right)} \phi(\{u_k\}, (x_0, \{w_k, v_k\})),$$

where, for fixed N, $\{w_k, v_k\}_{k=N}^{\infty} = 0$. To guarantee the problem is bounded, we require $\phi(\{u_k\}, \cdot)$ is concave for all $\{u_k\}$. This condition translates into rank condition, for $k = 0, \ldots, (N-1)$,

$$\gamma^2 I - G^T S_k G > 0,$$

where

$$
\begin{aligned}
S_N &= P \\
S_k &= Q + A^T\left(S_{k+1} + G^T S_{k+1}(\gamma^2 - G^T S_{k+1} G)^{-1} S_{k+1} G\right) A,
\end{aligned}
$$

and P denotes the solution of the algebraic Riccati equation:

$$
P = Q + A^T\left(P - B^T P(R + B^T P B)^{-1} P B\right) A.
$$

If we choose $\gamma > 0$ sufficiently large, then we can satisfy the rank condition for all N only if the state transition matrix A is has no eigenvalues on or outside the unit circle. We cannot expect a solution exists when the matrix A is unstable: the "adversary" can always choose a disturbance that trumps the open-loop control.

Example 9.3.1 *If we optimize over feedback policies (closed-loop control) rather than fixed controls (open-loop control), then the existence conditions are less restrictive. Consider the system*

$$
x_{k+1} = x_k + u_k + w_k
$$

and tuning parameters $Q = 1$, $R = 1$, and $P = 1$. The following table compares the minimum value of γ as a function of the horizon length N for open-loop and closed-loop control.

N	$\gamma_{\min}$ *Open-loop*	$\gamma_{\min}$ *Closed-loop*
5	*2.9*	*1.41*
10	*6.1*	*1.42*
15	*9.3*	*1.42*
50	*31.6*	*1.42*
∞	∞	*1.42*

The minimal value of γ sufficient for existence to a problem with a horizon N is roughly

$$
O\left(\sum_{k=1}^{N} k\lambda_{\max}(A)^k\right).
$$

The inability to have systems with unit eigenvalues (i.e. integrators) is especially restrictive from a practical standpoint because of the desire to add integral control. Our discussion of the infinite horizon regulator is for illustrative purposes, so we ignore the technical issues and tacitly assume existence. We return to issue of existence in Section 9.4.

Before demonstrating that the MPC achieves l_2-gain $\leq \gamma$, we introduce a forward-backward dynamic programming decomposition of $\mathcal{P}(T)$. An interesting discussion of forward-backward dynamic programming in an abstract setting can be found in (Verdu and Poor 1987). The details of the argument are given by Bernhard (1995). We begin by first introducing the **arrival cost** function. We say a disturbance sequence $\left(x_0, \{w_k, v_k\}_{k=0}^{T-1}\right)$ is **output admissible** at time T if it satisfies the state equation (9.1) subject to the (fixed) data $\{u_{k|k}^*, y_k\}_{k=0}^{T-1}$. Let $\mathbb{Q}_T$ denote the output admissible set of disturbances

$$
\mathbb{Q}_T = \left\{x_0, \{w_k, v_k\}_{k=0}^{T-1} \;:\; y_k = Cx\left(k; x_0, \{u_{j|j}^*\}, \{w_j\}\right) + Ev_k\right\}
$$

and

$$
Z_T(\{u_k\}, (x_0, \{w_k, v_k\})) := \frac{1}{2}\sum_{k=0}^{T-1} x_k^T Q x_k + u_k^T R u_k - \gamma^2\left(w_k^T w_k + v_k^T v_k + x_0^T \Pi^{-1} x_0\right).
$$

For all $z \in \mathbb{R}^n$, we define the arrival cost at time T as

$$\mathcal{Z}_T(z) := \max_{\left(x_0, \{w_k, v_k\}_{k=0}^{N-1}\right)} \left\{ Z_T(\{u_k\}, (x_0, \{w_k, v_k\})) \ : \ (x_0, \{w_k, v_k\}) \in \mathbb{Q}_T, \ \ x_T = z \right\},$$

where it is understood that $\mathcal{Z}_0(z) := -\gamma^2 z^T \Pi^{-1} z$. Arrival cost $\mathcal{Z}_T(x)$ is defined for all $x \in \mathbb{R}^n$, because of the form of the observation equation: $y_k = Cx_k + v_k$. Note that the input constraints $u_k \in \mathbb{U}$ are not present in $\mathcal{Z}_T(\cdot)$. We justify their omission on the following grounds: only the controls need to satisfy the constraints, and the controls are fixed for times less than T.

Likewise, we define the **cost to go** as

$$\mathcal{V}_T(z) = \min_{\{u_k\}_{k=0}^{\infty}} \max_{\{w_k, v_k\}_{k=0}^{\infty}} \left\{ V(\{u_k\}, \{w_k\}) \ ; \ x_k := x(k; z, \{u_j\}, \{w_j\}) \right\}, \tag{9.3}$$

where

$$V(\{u_k\}, \{w_j\}) := \frac{1}{2} \sum_{k=0}^{\infty} x_k^T Q x_k + u_k^T R u_k - \gamma^2 w_k^T w_k.$$

We ignore the effect of the output disturbances $\{v_k\}$ in $\mathcal{V}_T(z)$ as they do not affect the state directly; the output disturbances arise only in the estimation part of the control problem, not in regulation. So, we have the equivalence

$$\begin{aligned} V(\{u_k\}, \{w_j\}) &\equiv V(\{u_k\}, \{w_k, v_k\}), \\ &:= \frac{1}{2} \sum_{k=0}^{\infty} x_k^T Q x_k + u_k^T R u_k - \gamma^2 \left(w_k^T w_k + v_k^T v_k \right). \end{aligned}$$

Using the result of Bernhard (1995), we know that if

$$\hat{x}_T := \arg\max_z \mathcal{Z}_T(z) + \mathcal{V}_T(z), \tag{9.4}$$

where $\hat{x}_T$ denotes the $\mathcal{H}_\infty$ state estimate given the data $\{y_k\}_{k=0}^{T-1}$, then, assuming the problems (9.4) and (9.3) have unique solutions (exact conditions are stated later), the sequence

$$\{u_{k|T}^*\}_{k=T}^{\infty}$$

solves $\mathcal{P}(\hat{x}_T, T)$ and vice-versa. In particular,

$$\phi_T^* = \max_z \mathcal{Z}_T(z) + \mathcal{V}_T(z).$$

This result is significant, because it allows us to establish **certainty equivalence**.

When we implement MPC, we use, for computational purposes, a separation principle rather than certainty equivalence. In many ways, we can view the separation principle as the information state approach of James and Baras (1995). By application of forward dynamic programming (see Appendix C), we know

$$\mathcal{Z}_T(z) = -\frac{1}{2} \gamma^2 (z - \bar{x}_T)^T \Pi_T^{-1} (z - \bar{x}_T) + \alpha$$

where α is a constant,

$$\Pi_{k+1} = A \left(\Pi_k^{-1} + C^T (EE^T)^{-1} C - \gamma^{-2} Q \right) A^T + GG^T,$$

subject to the initial condition $\Pi_0 = \Pi$, and

$$\tilde{x}_{k+1} = A\tilde{x}_k + Bu_k^* + A \left(\Pi_k^{-1} + C^T (EE^T)^{-1} C - \gamma^{-2} Q \right) \left(\gamma^{-2} Q \tilde{x}_k + C^T (EE^T)^{-1} (y_k - C\tilde{x}_k) \right)$$

with $\tilde{x}_0 := 0$. As we prove in Appendix C,

$$\tilde{x}_T = \max_z \mathcal{Z}_T(z).$$

To obtain a separation, we can reformulate $\mathcal{P}(T)$ as the following dynamic game:

$$\begin{aligned} \phi_T^* &= \max_z \mathcal{V}_T(z) - \gamma^2 (z - \tilde{x}_T)^T \Pi_T^{-1} (z - \tilde{x}_T) + \alpha, \\ &= \min_{\{u_k\}_{k=0}^\infty} \max_{\left(x_0, \{w_k\}_{k=0}^\infty\right)} \phi_\infty(\{u_k\}, \left(x_0, \{w_k\}_{k=0}^\infty\right), \tilde{x}_T) + \alpha, \end{aligned}$$

and

$$\phi_T(\{u_k\}, \left(x_0, \{w_k\}_{k=0}^\infty\right), z) := \sum_{k=0}^\infty x_k^T Q x_k + u_k^T R u_k - \gamma^2 \left(w_k^T w_k + (x_0 - z)^T \Pi_T^{-1} (x_0 - z)\right).$$

To demonstrate infinite horizon MPC achieves finite l_2-gain $\leq \gamma$ closed-loop performance, consider time $T = 0$ and suppose the solution $\{u_{k|0}^*\}_{k=0}^\infty$ and $\left(x_0^*, \{w_{k|0}^*, v_{k|0}^*\}_{k=0}^\infty\right)$ exists. Because the solution $\{u_{k|j}^*\}_{k=0}^\infty$ and $\left(x_0, \{w_{k|j}^*, v_{k|j}^*\}_{k=0}^\infty\right)$ at time $j > 0$ is feasible at time $T = 0$, we know

$$\begin{aligned} \phi_0^* = \phi(\{u_{k|0}^*\}, \left(x_0, \{w_{k|0}^*, v_{k|0}^*\}\right)) &\geq \phi(\{u_{k|0}^*\}. \left(x_0, \{w_{k|j}^*, v_{k|j}^*\}\right)), \\ &\geq \phi(\{u_{k|j}^*\}. \left(x_0, \{w_{k|j}^*, v_{k|j}^*\}\right)), \\ &= \phi_j^*. \end{aligned}$$

The first inequality results from the addition of the history constraint in $\mathcal{P}(T)$ and second inequality results from optimality. We may view the second inequality as our dissipation inequality. Using the dynamic programming decomposition

$$\phi_j^* = \max_z \mathcal{Z}_j(z) + \mathcal{V}_j(z),$$

we have by optimality that, for all $j \geq 0$,

$$\begin{aligned} \max_z \mathcal{Z}_j(z) &\leq \phi_j^*, \\ &\leq \phi_0^*, \\ &\leq 0. \end{aligned}$$

The last inequality is proved in the Appendix C. From the above inequalities, we know that for all disturbances $\left(x_0, \{w_j, v_j\}_{j=0}^{k-1}\right) \in \mathbb{Q}_T$

$$\frac{1}{2} \sum_{k=0}^{T-1} x_k^T Q x_k + u_k^T R u_k - \gamma^2 \left(w_k^T w_k + v_k^T v_k + x_0^T \Pi^{-1} x_0\right) \leq 0,$$

or that

$$\begin{aligned} &\sum_{k=0}^{T-1} x_k^T Q x_k + u_k^T R u_k \leq \\ &\qquad \gamma^2 \left(\sum_{k=0}^{T-1} w_k^T w_k + v_k^T v_k + x_0^T \Pi^{-1} x_0 \right). \end{aligned}$$

Because we have limited our discussion to stable systems, infinite-horizon MPC achieves finite l_2-gain $\leq \gamma$ for all $(x_0, \{v_k, w_k\}_{k=0}^\infty)$. In other words, the admissible class of disturbances are all square summable sequences, i.e. $\mathcal{D}(\gamma) = \mathbb{R}^n \times l_2(\mathbb{R}^{q+p})$.

9.4 Finite-Horizon Formulation

In the preceding section we discussed the infinite-horizon formulation of constrained MPC. In addition to computational difficulties, application of the infinite horizon formulation is limited to systems with eigenvalues strictly inside the unit circle. To overcome these difficulties, we need to consider a finite-horizon formulation. Consider a finite-horizon game with an open-loop information structure of the form

$$\mathcal{P}_N(x,T,\gamma): \qquad \phi_N^*(x,T) = \min_{\pi_N \in \mathbb{U}^N} \max_{\left(x_0, \{w_k\}_{k=0}^{N-1}\right)} \phi_N(\{u_k\},(x_0,\{w_k\})\,;x,T,\gamma),$$

subject to the system

$$x_{k+1} = Ax_k + Bu_k + Gw_k,$$

where the objective function is defined by

$$\Phi_N(\{u_k\},(x_0,\{w_k\})\,;x,T,\gamma) = \\ \sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k + x_N^T P x_N - \gamma^2\left(w_k^T w_k + (x_0-x)^T \Pi_T^{-1}(x_0-x)\right)$$

and the matrices Q, R, P, and Π_T are positive definite.

To guarantee a solution exists to the dynamic game $\mathcal{P}_N(x,T,\gamma)$, we need to choose γ sufficiently large such that the following rank conditions are satisfied (Başar and Olsder 1999),

$$\gamma^2 I - G^T S_k G > 0, \qquad k = 0,\dots,N-1$$

and

$$\lambda_{\max}(\Pi_T S_0) < \gamma^2$$

where

$$\begin{aligned} S_N &= P, \\ S_k &= Q + A^T(S_{k+1} + S_{k+1}G(\gamma^2 - G^T S_{k+1} G)^{-1} G^T S_{k+1})A. \end{aligned}$$

We establish the existence of γ in the following proposition.

Proposition 9.4.1 *If we assume $\mathbb{U}$ contains the origin, then, given $N > 0$ and $T \geq 0$, there exists $\bar{\gamma} \in \Gamma$ such that for all $\gamma > \bar{\gamma}$, $x \in \mathbb{R}^n$, a solution exists to $\mathcal{P}_N(x,T,\gamma)$.*

Proof. As the set $\mathbb{U}$ is non-empty, it suffices to establish the existence of a constant γ such that $\phi(\{u_k\},\cdot;\gamma)$ is strictly concave. If we write out the concavity condition explicitly, then this condition translates into the following matrix inequality:

$$\gamma^2 \mathbb{Q}^{-1} - \mathbb{G}\mathbb{P}\mathbb{G}^T > 0, \tag{9.5}$$

where

$$\mathbb{Q} = \operatorname{diag}\left(\begin{bmatrix} \Pi_T & I & \cdots & I \end{bmatrix}\right),$$

and

$$\mathbb{G} = \begin{bmatrix} A & G & & & \\ A^2 & AG & G & & \\ \vdots & \vdots & \vdots & \ddots & \\ A^N & A^{N-1}G & A^{N-2}G & \cdots & G \end{bmatrix}, \qquad \mathbb{P} = \begin{bmatrix} P & & & \\ & Q & & \\ & & \ddots & \\ & & & Q \end{bmatrix}.$$

If we choose $\gamma > \lambda_{\max}(\mathbb{Q}\mathbb{G}\mathbb{P}\mathbb{G}^T) \vee \lambda_{\max}(\Pi_T S_0)$, then the matrix inequality(9.5) is satisfied. $\square$

Remark 9.4.2 *If we assume the pair (A, C) is observable, then there exists a minimal positive definite solution P_∞ to the generalized algebraic Riccati equation*

$$\Pi_\infty = A\left(\Pi_\infty^{-1} + C^T(EE^T)^{-1}C - \gamma^{-2}Q\right)A^T + GG^T.$$

If we choose $\Pi = \Pi_\infty$, then $\Pi_k = \Pi_\infty$ for all $k \geq 0$. In this case, there exists a $\gamma > 0$ such that a solution exists to $\mathcal{P}_N(x, T, \gamma)$ for all $x \in \mathbb{R}^N$ and $T \geq 0$.

Remark 9.4.3 *Suppose $\gamma > 0$ is sufficiently large such that a solution exists to $\mathcal{P}_N(x, T, \gamma)$ and consider the following compact notational representation of the dynamic game $\mathcal{P}_N(x, T, \gamma)$*

$$\mathcal{P}_N(x, T, \gamma): \qquad \min_{\mathbf{U} \in \mathbb{U}^N} \max_{\mathbf{W}} \mathbf{x}^T \mathcal{Q}\mathbf{x} + \mathbf{u}^T \mathcal{R}\mathbf{u} - \mathbf{w}^T \Gamma \mathbf{w}$$

subject to constraint

$$\mathbf{x} = \mathbb{A}x + \mathbb{B}\mathbf{u} + \mathbb{G}\mathbf{w}.$$

The max *is unconstrained, so we can solve for $\mathbf{w}$ analytically, and, consequently, we are able to reformulate the dynamic game $\mathcal{P}_N(x, T, \gamma)$ as the following quadratic program*

$$\mathcal{P}_N(x, T, \gamma): \qquad \min_{U \in \mathbb{U}^N} \pi_N^T \mathcal{H} \pi_N + \mathcal{C}^T \pi_N$$

where

$$\begin{aligned}
\mathcal{H} &= \mathcal{R} + \mathbb{B}^T\left(\mathcal{Q} + \mathcal{Q}\mathbb{G}\left(\Gamma - \mathbb{G}^T\mathcal{Q}\mathbb{G}\right)^{-1}\mathbb{G}^T\mathcal{Q}\right)\mathbb{B}, \\
\mathcal{C} &= \mathbb{B}^T\left(\mathcal{Q} + \mathcal{Q}\mathbb{G}\left(\Gamma - \mathbb{G}^T\mathcal{Q}\mathbb{G}\right)^{-1}\mathbb{G}^T\mathcal{Q}\right)\mathbb{A}x.
\end{aligned}$$

The cost, therefore, to "robustify" MPC is negligible.

9.4.1 The Terminal Cost and Constraint Set

For some $\gamma > 0$, suppose $P(\gamma)$ is the minimal positive definite solution of the generalized algebraic Riccati equation

$$P(\gamma) = Q + A^T(P^{-1}(\gamma) + BR^{-1}B^T - \gamma^{-2}GG^T)^{-1}A \tag{9.6}$$

satisfying the following conditions

A : $\lambda_{\max}(GP(\gamma)G^T) < \gamma^2$;

B : $Q + A^T\Lambda^{-T}P(\gamma)\left(BR^{-1}B^T - \gamma^{-2}GG^T\right)P(\gamma)\Lambda^{-1} \geq 0$,

where

$$\Lambda = I + (BR^{-1}B^T - \gamma^{-2}GG^T)P(\gamma).$$

Under the stated conditions, condition **A** is sufficient to guarantee the matrix $P(\gamma)$ solves the infinite horizon game in that it generates the minimal value function for the stationary Isaacs equation (Başar 1991)

$$x^T P(\gamma)x = \min_u \max_w \left\{x^T Qx + u^T Ru - \gamma^2 w_k^T w_k + (Ax + Bu + Gw)^T P(\gamma)(Ax + Bu + Gw)\right\}.$$

The optimal solution is

$$\begin{aligned} u^* &= -R^{-1/2}B^T P(\gamma)\Lambda^{-1}Ax, \\ &:= K_u x, \\ w^* &= \gamma^{-2}G^T P(\gamma)\Lambda^{-1}Ax, \\ &:= K_w x. \end{aligned}$$

The feedback law K_u is the state feedback $\mathcal{H}_\infty$ controller (Doyle, Glover, Khargonekar and Francis 1989). As the robustness margins increase as γ decreases, one typically calculates K_u iteratively by reducing γ until condition **A** is violated.

Condition **B** is necessary to guarantee that the matrix $P(\gamma)$ satisfies the following Lyapunov inequality

$$P(\gamma) - A\Lambda^{-T}P(\gamma)\Lambda^{-1}A \geq 0, \tag{9.7}$$

which is precisely a restatement of condition **B**. Why is the inequality (9.7) useful? The answer becomes evident when we consider the consequences of the following proposition.

Proposition 9.4.4 *Suppose $P(\gamma)$ is a minimal positive definite solution of (9.6) satisfying conditions* **A** *and* **B**. *If $x^T P(\gamma)x \leq \alpha$, then $x_+^T P(\gamma)x_+ \leq \alpha$, where $x_+ = (A + BK_u + GK_w)x = \Lambda^{-1}Ax$.*

Proof. As the matrix $P(\gamma)$ is positive definite, it suffices to demonstrate

$$x^T Qx + u^{*T}Ru^* - \gamma^2 w^{*T}Qw^* \geq 0$$

where the vectors u^* and w^* denote the optimal solution to the infinite horizon game. In particular, we know $u^* = K_u x$ and $w^* = K_w w$. If substitute in K_u and K_w, then one obtains condition **B**, and the proposition follows as claimed. □

Remark 9.4.5 *Condition* **B** *is equivalent to stating that the matrix $P(\gamma)$ generates a (stable) Lyapunov function for the system $x_{k+1} = \Lambda^{-1}Ax_k$. While this system is stable (Başar and Bernhard 1995), it does not appear* **B** *is automatically satisfied. Condition* **B** *is satisfied when $BR^{-1}B^T - \gamma^{-2}GG^T > 0$. Obviously, this condition is very restrictive. In the context of developing an invariant set, we can replace the level set with a Gilbert and Tan set (Gilbert and Tan 1991) as the system $x_{k+1} = \Lambda^{-1}Ax_k$ is stable (Başar and Bernhard 1995). Consequently, condition* **B** *is not necessary.*

A set $\mathbb{Z} \subset \mathbb{R}^n$ is **positive invariant** for the dynamic system Σ, if for every initial state $x_0 \in \mathbb{Z}$, the subsequent motion x_k, $k \geq 0$, belongs to $\mathbb{Z}$. A set $\mathbb{Z} \subset \mathbb{R}^n$ is $\mathbb{Y}$ **output admissible** for the dynamic system Σ, where $\mathbb{Y} \subset \mathbb{R}^n$, if for every initial state $x_0 \in \mathbb{Z}$, the subsequent motion x_k, $k \geq 0$, belongs to $\mathbb{Y}$.

Consider the level set

$$L_\alpha = \left\{x \;:\; x^T P(\gamma)x \leq \alpha\right\},$$

where $\alpha > 0$. Condition **A** and **B**, in light of Proposition 9.4.4, guarantee the level set L_α is positively invariant for the dynamic system $x_{k+1} = \Lambda^{-1}Ax_k$. Consider now the constraint set

$$\mathbb{X} = \{x \;:\; K_u x \in \mathbb{U}\}.$$

If $0 \in \operatorname{int} \mathbb{X}$, then there exists $\epsilon > 0$ such $B_\epsilon \subset \mathbb{X}$. If we choose

$$\alpha = \frac{\epsilon^2}{\lambda_{\max}(P(\gamma))},$$

then $L_\alpha \subset \mathbb{U}$ and the level set L_α is $\mathbb{X}$ output admissible. One may also construct the maximal output admissible set using the Gilbert and Tan (1991) algorithm (see Remark 9.4.5).

The invariant set is important, because it allows us to establish MPC has finite l_2-gain$\leq \gamma$. In particular, if $x(N; x_0^*\{u_j^*\}, \{w_j^*\}) \in L_\alpha \subset \mathbb{U}$, then $\mathcal{P}_N(\cdot) = \mathcal{P}_\infty(\cdot)$. We are able to bypass the existence issue, because the input sequence $\{u_k\}_{k=N}^{\infty}$ is now replaced with a sequence of feedback policies $\{\mu_k(\cdot) = K_u(\cdot)\}_{k=N}^{\infty}$. One can demonstrate there exists a $\gamma > 0$ such that $x(N; x_0^*\{u_j^*\}, \{w_j^*\}) \in L_\alpha \subset \mathbb{U}$. However, there is one significant caveat: invariance as we have constructed it is with respect to the worst cast disturbance $w_k = K_w x_k$. Our construction immediately begs the question "what else are we invariant to." We address this question in the next section.

9.5 Constrained reachability of perturbed systems

Let $\gamma > 0$. We say the level set $\mathcal{L}_\alpha$ is γ-**strongly reachable** at time N from the origin if there exists an input sequence $\{u_k\}_{k=0}^{N-1} \in \mathbb{U}^N$ such that $x(N; x_0, \{u_j\}, \{w_j\}) \in \mathcal{L}_\alpha$ for all $\left(x_0, \{w_k\}_{k=0}^{N-1}\right) \in \mathcal{D}_N(\gamma)$, where the set of admissible disturbances $\mathcal{D}_N(\gamma)$ is an ellipsoid of the form

$$\mathcal{D}_N(\gamma) = \left\{ \left(x_0, \{w_k\}_{k=0}^{N}\right) \; : \; \sum_{k=0}^{N-1} w_k^T w_k + x_0 \Pi^{-1} x_0 \leq \frac{\alpha}{\gamma^2} \right\},$$

The reader is directed to Delfour and Mitter (1969) for a general discussion of the open-loop reachability problem.

Consider the following dynamic game of the pursuit-evasion class

$$\mathcal{R}_N(\gamma): \qquad \theta_N(\gamma) = \min_{\{u_k\}_{k=0}^{N-1} \in \mathbb{U}^N} \max_{\{w_k\}_{k=0}^{N-1}} x_N^T P x_N - \gamma^2 \sum_{k=0}^{N} w_k^T w_k,$$

where $x_N := x(N; z, \{u_j\}, \{w_j\})$. Let $\{u_k^*\}$ denote the solution to $\mathcal{R}_N(\gamma)$. For a solution to exist to $\mathcal{R}_N(\gamma)$ (Başar and Olsder 1999), we require that γ satisfies the following matrix inequality

$$\gamma^2 I > G^T S_k G, \qquad k = 0, \ldots, (N-1) \tag{9.8}$$

where

$$\begin{aligned} S_N &= P, \\ S_{N-1} &= A^T \left(S_N - S_N G (\gamma^2 - G^T S_N G)^{-1} G^T S_N\right) A. \end{aligned}$$

Our claim is that if a solution exists to $\mathcal{R}_N(\gamma)$ for some $\gamma > 0$, then the set $\mathcal{L}_\alpha$ is γ-**strongly reachable** at time N.

Proposition 9.5.1 *Suppose the pair (γ, N) guarantee a solution exists to $\mathcal{R}_N(\gamma)$. Then, there exists an input sequence $\{u_k^*\}_{k=0}^{N-1}$ solving the dynamic game $\mathcal{R}_N(\gamma)$ such that $x(N; z, \{u_j^*\}, \{w_j\}) \in \mathcal{L}_\alpha$ for all $(x_0, \{w_j\}) \in \mathcal{D}_N(\gamma)$ and the set $\mathcal{L}_\alpha$ is* γ-**strongly reachable** *at time N.*

Proof. Existence of a solution follows from the convexity and strict concavity of the game $\mathcal{R}_N(\gamma)$. Writing out the solution to the game explicitly, we have by optimality

$$\begin{aligned} \theta_N(\gamma) &= \max_{\left(x_0, \{w_k\}_{k=0}^{N-1}\right)} x(N; x_0, \{u_j^*\}, \{w_j\})^T P x(N; x_0, \{u_j^*\}, \{w_j\}) - \gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + x_0^T \Pi x_0 \right), \\ &\geq x(N; \bar{x}_0, \{u_j^*\}, \{\bar{w}_j\})^T P x(N; \bar{x}_0, \{u_j^*\}, \{\bar{w}_j\}) - \gamma^2 \left(\sum_{k=0}^{N-1} \bar{w}_k^T \bar{w}_k + \bar{x}_0^T \Pi \bar{x}_0 \right), \\ &\qquad \text{for all } \left(\bar{x}_0, \{\bar{w}_k\}_{k=0}^{N-1}\right). \end{aligned} \tag{9.9a}$$

Let

$$\beta = x(N; x_0^*, \{u_j^*\}, \{w_j^*\})^T P x(N; x_0^*, \{u_j^*\}, \{w_j^*\})$$

and $\delta = \alpha - \beta$, which is non-negative by assumption. Suppose the contrapositive and assume there exists some $\left(x_0, \{w_k\}_{k=0}^{N-1}\right) \in \mathcal{D}_N(z, \gamma)$ such that

$$x(N; x_0, \{u_j^*\}, \{w_j\})^T P x(N; x_0, \{u_j^*\}, \{w_j\}) > \alpha.$$

Let

$$\theta_N'(\gamma) = x(N; x_0, \{u_j^*\}, \{w_j\})^T P x(N; x_0, \{u_j^*\}, \{w_j\}) - \gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + x_0^T \Pi x_0 \right).$$

By assumption, we obtain the inequality

$$\theta_N'(\gamma) + \gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + x_0^T \Pi x_0 \right) > \theta_N(\gamma) + \gamma^2 \left(\sum_{k=0}^{N-1} w_k^{*T} w_k^* + x_0^{*T} \Pi x_0^* \right) + \delta = \alpha.$$

By the inequality (9.9a), it follows necessarily that $\theta_N(\gamma) \geq \theta_N'(\gamma)$. However, these two inequalities contradicts the main assumption: that is

$$\gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + x_0^T \Pi x_0 \right) > (\theta_N(\gamma) - \theta_N'(\gamma)) + \gamma^2 \left(\sum_{k=0}^{N-1} w_k^{*T} w_k^* + x_0^T \Pi x_0 \right) + \delta = \alpha,$$

and, therefore, we obtained the desired result. □

Proposition 9.5.1 is significant, because it characterizes the class of disturbances MPC can reject. As we prove in the next proposition, our characterization of the admissible class of disturbances is not conservative.

Proposition 9.5.2 *Suppose the pair (γ, N) guarantee a solution exists to $\mathcal{R}_N(\gamma)$ and let $\{u_k^*\}_{k=0}^{N-1}$ denote the solution to $\mathcal{R}_N(\gamma)$. Then, for all $\epsilon > 0$, there exists a sequence $(\bar{x}_0, \{\bar{w}_k\}_{k=0}^{N-1})$ such that*

$$\gamma^2 \left(\sum_{k=0}^{N-1} \bar{w}_k^T \bar{w}_k + \bar{x}_0^T \Pi \bar{x}_0 \right) \leq \alpha + \epsilon$$

and $x(N; \bar{x}_0, \{u_j^*\}, \{\bar{w}_j\}) \notin \mathcal{L}_\alpha$.

Proof. First note that $\{u_j^*\}_{k=0}^{N-1} = 0$. To prove this claim, suppose $\{u_k\}_{k=0}^{N-1} = 0$ and consider the following.

$$\begin{aligned} &- \max_{x_0, \{w_k\}_{k=0}^{N-1}} \left\{ -\gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + x_0^T \Pi x_0 \right) \ : \ x(N; x_0, \{u_j^*\}, \{w_j\})^T P x(N; x_0, \{0\}, \{w_j\}) = 0 \right\} \\ &= \min_{x_0, \{w_k\}_{k=0}^{N-1}} \left\{ \gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + x_0^T \Pi x_0 \right) \ : \ x(N; x_0, \{u_j^*\}, \{w_j\})^T P x(N; x_0, \{0\}, \{w_j\}) = 0 \right\} = 0. \end{aligned}$$

As the matrix P is positive definite, we cannot improve on the choice $\{u_k\} = \{0\}$, and $\{u_j^*\}_{k=0}^{N-1} = 0$ (though the solution may not be unique). Now let

$$c = \min_{x_0, \{w_k\}_{k=0}^{N-1}} \left\{ \gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + x_0^T \Pi x_0 \right) \ : \ x(N; x_0, \{u_j^*\}, \{w_j\})^T P x(N; x_0, \{u_j^*\}, \{w_j\}) = \alpha \right\}$$

and $\mathbf{u} = \left(x_0^0, \{w_k^0\}_{k=0}^{N-1}\right)$ denote the solution to the minimization. As the system (9.1) is linear, there exists a matrix L such that

$$x(N; x_0^0, \{u_j^*\}, \{w_j^0\}) = L\mathbf{w}.$$

and

$$\mathbf{w}^T L^T P L \mathbf{w} = \alpha.$$

If we choose $\mathbf{w}^0 = \delta\mathbf{w}$ for some $\delta > 1$, then

$$\mathbf{w}^{0T} L^T P L \mathbf{w}^0 = \delta^2 \alpha > \alpha.$$

From proposition 9.5.1, we know that $c \leq \alpha$. If we choose

$$\delta < \frac{c+\epsilon}{c},$$

then the proposition follows as claimed. □

If there are no constraints ($\alpha = \infty$ and $\mathbb{U} = \mathbb{R}^n$), then the set of admissible disturbances $\mathcal{D}_N(\gamma) = \mathbb{R}^{Nq}$ for all $\gamma > 0$ as expected. Without any limitations the controller can respond in equal magnitude to any disturbance. This distinguishes unconstrained from constrained control.

One heuristic for γ is that is scales roughly as $O(\sum_{k=1}^{N} k\lambda_{\max} A^k)$. One may view, therefore, γ as a bound on the open-loop dynamics of the system (9.1). Open-loop control, consequently, does not improve the admissible class of disturbances $\mathcal{D}_N(\gamma)$. One improve the "robustness" of a system only if feedback is employed. However, as a consequences of eschewing dynamic programming, the theory and application of MPC explicitly relies on the open-loop properties of control. The calculated margins of MPC are consequently the same as pure open-loop control. We discuss these issues further in the conclusion.

9.6 Admissible Class of Disturbances

In the preceding section we characterized the admissible class of disturbances for the simple problem of driving the state into the level set $\mathcal{L}_\alpha$. In this section, we extend those results to MPC.

Consider again the dynamic game

$$\mathcal{P}_N(x,T,\gamma): \qquad \phi_N^*(x,T) = \min_{\{u_k\}_{k=0}^{N-1} \in \mathbb{U}^N} \ \max_{\left(x_0, \{w_k, v_k\}_{k=0}^{N-1}\right)} \phi_N(\{u_k\}, ((x_0, \{w_k, v_k\})\,; T, \gamma)$$

Suppose N and γ are chosen such that a solution exists. One choice for the admissible class of disturbances is then

$$\mathcal{D}(\gamma) = \left\{ (x_0, \{w_k\}_{k=0}^{\infty}) \ : \ \sum_{k=0}^{\infty} w_k^T w_k + v_k^T v_k + x_0^T \Pi^{-1} x_0 \leq \frac{\alpha}{\gamma^2} \right\}$$

The notation $\left(x_0, \{w_k\}_{k=0}^{N-1}\right) \in \mathcal{D}(\gamma)$ is used to denote $(x_0, \{w_0, \ldots, w_{N-1}, 0, 0, \ldots\}) \in \mathcal{D}(\gamma)$.

Proposition 9.6.1 *Suppose the pair (γ, N) guarantees a solution exists to $\mathcal{P}_N(x,T,\gamma)$ for all $x \in \mathbb{R}^n$. Then, the state $x(N; x_0, \{u_j^*\}, \{w_j\}) \in L_\alpha$ for all $\left(x_0, \{w_k\}_{k=0}^{N-1}\right) \in \mathbb{D}(\gamma)$ where $\{u_j^*\}$ denotes the solution to $\mathcal{P}_N(0,\gamma)$.*

Proof. The proof is omitted due to the close similarity with the proof of Proposition 9.5.1. □

Let $\{u_{j|k}\}_{k=0}^{N-1}$ denote the solution, assuming it exists, of $\mathcal{P}_N(x,j,\gamma)$ and $u_k^0 := u_{0|k}^*$.

Proposition 9.6.2 *Suppose the pair (γ, N) guarantee a solution exists to $\mathcal{P}_N(x,T,\gamma)$ for all $x \in \mathbb{R}^n$ and $T \geq 0$ and the matrix $P(\gamma)$ is the minimal solution of (9.6) satisfying conditions* **A** *and* **B**. *If $\mathcal{L}_\alpha \subset \mathbb{U}$, then*

$$\phi_N^*(0,0) \geq \phi_N^*(\tilde{x}_k, k)$$

for all $\left(x_0, \{w_k\}_{k=0}^N\right) \in \mathcal{D}(\gamma)$ and the closed-loop system has finite l_2-gain$\leq \gamma$.

Proof. Let $(\bar{z}, \{\bar{w}_k\}_{k=0}^\infty \in \mathcal{D}(\gamma)$. We proceed using an induction argument. Using the properties of the matrix $P(\gamma)$, we have

$$\begin{aligned}
\phi_N(0,\gamma) \;\; &= \;\; \min_{\{u_k\}_{k=0}^{N-1}\in\mathbb{U}^N} \max_{\left(x_0,\{w_k\}_{k=0}^{N-1}\right)} \sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k + x_N^T P(\gamma) x_N \\
&\quad -\gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + v_k^T v_k + x_0^T \Pi^{-1} x_0 \right) \\
&\geq \;\; \min_{\{u_k\}_{k=0}^{N-1}\mathbb{U}^N} \max_{\left(x_0,\{w_k\}_{k=0}^{N-1}\right)} \min_{u_N\in\mathbb{U}} \max_{w_N} \; u_N^T R u_N + x_N^T Q x_N + x_{N+1}^T P(\gamma) x_{N+1} - \gamma^2 w_N^T w_N + \\
&\quad \sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k - \gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + v_k^T v_k + x_0^T \Pi^{-1} x_0 \right) \\
&\quad (\text{as } x_N^T P(\gamma) x_N \leq \alpha) \\
&\geq \;\; \min_{\{u_k\}_{k=0}^{N}\in\mathbb{U}^N} \max_{\left(x_0,\{w_k\}_{k=0}^{N}\right)} \sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k + x_{N+1}^T P(\gamma) x_{N+1} - \\
&\qquad\qquad \gamma^2 \left(\sum_{k=0}^{N} w_k^T w_k + v_k^T v_k + x_0 \Pi^{-1} x_0 \right), \\
&\quad \left(\text{as } \min_x \max_y f(x,y) \geq \max_y \min_x f(x,y) \right) \\
&\geq \;\; \min_{\{u_k\}_{k=1}^{N}\in\mathbb{U}^N} \max_{\left(x_0,\{w_k\}_{k=0}^{N}\right)} \left\{ \sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k + x_{N+1}^T P(\gamma) x_{N+1} - \right. \\
&\qquad \left. \gamma^2 \left(\sum_{k=0}^{N} w_k^T w_k + v_k^T v_k + x_0 \Pi^{-1} x_0 \right) \; : \; u_0 = u_0^* \right\}, \\
&\geq \;\; \min_{\{u_k\}_{k=1}^{N}\in\mathbb{U}^N} \max_{\left(\{w_k\}_{k=1}^{N}\right)} \left\{ \sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k + x_{N+1}^T P(\gamma) x_{N+1} - \right. \\
&\qquad \left. \gamma^2 \left(\sum_{k=0}^{N} w_k^T w_k + v_k^T v_k + x_0 \Pi^{-1} x_0 \right) \; : \; u_0 = u_0^*, y_0 = C\bar{z} + \bar{v}_0 \right\}, \\
&= \;\; \phi_N(\tilde{x}_1, \gamma), \\
&\quad (\text{by the properties of } \Pi_1).
\end{aligned}$$

Let us assume $\phi_N(\tilde{x}_j, j-1) \geq \phi_N(\tilde{x}_j, j)$ and consider $\phi_N(\hat{x}_j, j)$.

$$\begin{aligned}
\phi_N^*(\tilde{x}_j,\gamma) \quad &= \quad \min_{\{u_k\}_{k=j}^{N+j-1} \in \mathbb{U}^N} \; \max_{\left(x_0, \{w_k\}_{k=j}^{N+j-1}\right)} \sum_{k=j}^{N+j-1} x_k^T Q x_k + u_k^T R u_k + x_N^T P(\gamma) x_N \\
&\quad -\gamma^2 \left(\sum_{k=0}^{N-1} w_k^T w_k + v_k^T v_k + (x_j - \hat{x}_j)^T \Pi_j^{-1} (x_j - \hat{x}_j) \right) \\
&\geq \quad \min_{\{u_k\}_{k=j}^{N+j-1} \in \mathbb{U}^N} \; \max_{\left(x_0, \{w_k, v_k\}_{k=0}^{N+j-1}\right)} \left\{ \sum_{k=j}^{N+j-1} x_k^T Q x_k + u_k^T R u_k + x_N^T P(\gamma) x_N \right. \\
&\quad -\gamma^2 \left(\sum_{k=0}^{N+j-1} w_k^T w_k + v_k^T v_k + x_0^T \Pi^{-1} x_0 \right) - \alpha_j \; : \\
&\quad \left. \begin{array}{c} \{u_k\}_{k=0}^{j-1} = \{u_k^0\}_{k=0}^{j-1} \\ \{y_k\}_{k=0}^{j-1} = \{Cx(k; \bar{z}, \{u_l^0\}, \{\bar{w}_l\}) + \bar{v}_k\}_{k=0}^{j-1} \end{array} \right\}.
\end{aligned}$$

If we let $\{u_k^j\}_{k=0}^{N-1}$ and $\left(z^j, \{w_k^j\}_{k=0}^{N-1}\right)$ denote the solution of $\mathcal{P}_N(\tilde{x}_j, j, \gamma)$, then it follows from Propositions 9.6.1 and 9.6.1 that

$$x(N; z^j, \{u_l^j\}, \{w_l^j\}) \in \mathcal{L}_\alpha.$$

Therefore, it follows from the above arguments that $\phi_N(\tilde{x}_j, j) \geq \phi_N(\tilde{x}_{j+1}, j+1)$ and the proposition follows as claimed. $\square$

As with the reachability problem, one may view the admissible class of disturbances as the set of disturbances such that the free evolution of the system does not leave level set $\mathcal{L}_\alpha$ at time N. The remarkable fact of this statement is that nowhere does feedback enter the problem. The *calculated* stability margins are solely open-loop margins. We emphasize that these are theoretical margins; the actual margin are much better. We elaborate on this idea in the Conclusion.

9.6.1 Remarks

When we implement MPC online (see Chapter 6), we typically vary the horizon length such that

$$x(N; \hat{x}, \{u_j^*\}, \{0\}) \in \mathcal{O}_\infty,$$

where $\hat{x}$ denotes our current measurement of the state and $\{u_k^*\}_{k=0}^{N-1}$ denotes the open-loop optimal control. It is possible to establish the existence of an N and γ such that $x(N; \hat{x}, \{u_i^*\}, \{w_j^*\}) \in \mathcal{O}_\infty$. However, this result is meaningless, because the horizon length N, attenuation level γ, and the initial condition $\hat{x}$ are intimately related. If we increase the horizon length N, then γ increases and the set of admissible disturbances shrinks. Furthermore, γ is a function of the initial condition $\hat{x}$. If the initial condition is relatively small (i.e. $||\hat{x}||$ small), then γ, and admissible class of disturbances, is relatively large. If the initial condition is relatively large (i.e. $||\hat{x}||$ large, then γ relatively small. So, loosely speaking, the controller is robust only if it is idle. As we have emphasized, this is a consequence of open-loop feedback control. Consequently, when we implement "robust" MPC, we need to fix N and γ. Otherwise, we are unable to quantify stability margins *a priori*.

9.7 Summary

For fixed N and γ, we formulate ***robust*** MPC as the following open-loop finite-horizon dynamic game

$$\mathcal{P}_N(x,T,\gamma): \qquad \phi_N^*(x,T) = \min_{\pi_N \in \mathbb{U}^N} \max_{\left(x_0,\{w_k\}_{k=0}^{N-1}\right)} \phi_N(\{u_k\},(x_0,\{w_k\})\,;x,T,\gamma),$$

where the objective function is defined by

$$\Phi_N(\{u_k\},(x_0,\{w_k\})\,;x,T,\gamma) = \sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k + x_N^T P(\gamma) x_N - \gamma^2 \left(w_k^T w_k + (x_0 - \tilde{x}_T)^T \Pi_T^{-1} (x_0 - \tilde{x}_T)\right)$$

and the estimator is defined by the dynamic system

$$\tilde{x}_{k+1} = A\tilde{x}_k + Bu_k^* + A\left(\Pi_k^{-1} + C^T(EE^T)^{-1}C - \gamma^{-2}Q\right)\left(\gamma^{-2}Q\tilde{x}_k + C^T(EE^T)^{-1}(y_k - C\tilde{x}_k)\right)$$

where

$$\Pi_{k+1} = A\left(\Pi_k^{-1} + C^T(EE^T)^{-1}C - \gamma^{-2}Q\right)A^T + GG^T,$$

subject to the initial conditions $\Pi_0 = \Pi$ and $\tilde{x}_0 = 0$. If $\{u_k(\tilde{x}_T)^*\}_{k=0}^{N-1}$ denote the solution to $\mathcal{P}_N(\tilde{x}_T,T,\gamma)$, then the feedback law is defined as

$$\mu_T(y_0,\ldots,y_{T-1},\mu_0,\ldots,\mu_{T-1}) := u_0^*(\tilde{x}_T).$$

If Π_k satisfies the matrix inequality

$$\gamma^2\Pi_k - Q > 0 \qquad k \geq 0,$$

and $P(\gamma)$ satisfies conditions **A** and **B** and the matrix inequalities

$$\gamma^2 I - G^T S_k G > 0, \qquad k = 0,\ldots,N-1,$$

and

$$\lambda_{\max}(\Pi_T S_0) < \gamma^2,$$

where

$$\begin{aligned} S_N &= P(\gamma), \\ S_k &= Q + A^T(S_{k+1} + S_{k+1}G(\gamma^2 - G^T S_{k+1}G)^{-1}G^T S_{k+1})A, \end{aligned}$$

then closed-loop system has finite l_2-gain$\leq \gamma$ for the class of disturbances $\mathcal{D}(\gamma)$.

9.8 Conclusion

In this chapter we formulated MPC as a dynamic game. The game formulation allowed us to obtain a separation for output feedback and prove that there exists a fixed $\gamma > 0$ such that the closed-loop system has finite l_2-gain$\leq \gamma$. Furthermore, the added cost associated with formulating MPC as a dynamic game is negligible; the resulting problem is a quadratic program, though the optimization problem is no longer sparse.

However, when we characterize the admissible class of disturbances, we are confronted with an apparent paradox: the closed-loop stability margins are equivalent to the open-loop margins. The game formulation of MPC does not improve robustness. These results are not surprising; robustness requires feedback (c.f. (Mayne et al. 2000)). By eschewing dynamic programming, we are unable to quantify the feedback character of MPC. The only statement we are able to make is that MPC performs no worse than open-loop control (c.f. (Bertsekas 1972)). This statement is the basis for the stability results derived in Chapter 5. The initial open-loop calculation by construction drives the state into terminal constraint set $\mathcal{X}_f$. Once the state is in the terminal constraint set $\mathcal{X}_f$, the local feedback policy $\kappa_f(\cdot)$ stabilizes the system. When we prove stability, we demonstrate only that the subsequent open-loop calculations do not decrease performance. No where in the proof do we demonstrate that feedback improves performance; rather, we show it does not degrade performance. These limitations are inherent in all minimax formulations of MPC. The most one can achieve with a minimax formulation is risk-sensitive open-loop control; e.g. operate the reactor away from its ignition temperature (c.f. (Ray and Barney 1972) and (Abel and Marquardt 1998)).

MPC is a feedback policy, and practice indicates it has good robustness properties. However, the goal in robust control is to design a controller with known stability/robustness margins. As MPC is a nonlinear control law, one quantifies the margins by including them as constraints in the design. This procedure inevitably leads to a minimax formulation. Unless one is willing to solve a Hamilton-Jacobi-Bellman-Isaacs inequality (Lin and Byrnes 1996), a computationally prohibitive task, it appears impossible to quantify the robustness properties of MPC. Alternatives using LMI's have recently been proposed (Kothare et al. 1996), though these approaches mimic reference governor strategies (c.f. (Bemporad and Mosca 1998)) and no longer explicitly use a prediction horizon (one of the appealing features of MPC). However, they optimize over linear feedback policies and the resulting properties accrue. It is not clear with these methods what one gains in robustness or loses in performance. How to satisfactorily resolve the issue of robustness in MPC is an open research problem.

One proposed solution is to parameterize the control with a linear feedback law (Chen et al. 1997). By precompensating with linear feedback, the system dynamics are stabilized, thereby eliminating the restrictive existence conditions. The cost is that precompensation introduces state constraints. In our development we explicitly avoided state constraints. The reason is that state constraints introduce the possibility of infeasibility: i.e. there is no solution to the minimization. As a possible outcome in the game, we need to quantify infeasibility as a choice for the adversary. One simply cannot just add constraints to the minimization - extra measures are necessary. One possibility is replace the state constraints with log-barrier penalty functions. Otherwise, the minimax formulation is not properly hedging against the worst-case scenario.

While margins are important, they are not critical for implementation. Most feedback controllers are not robustly designed, yet their performance is excellent. Furthermore, robustly designed controllers tend to perform poorly in practice. If robustness is an issue, the author suggests one use "ad-hoc" techniques such as integral control or loop transfer recovery (c.f. (Doyle and Stein 1979)). One may view these "ad hoc" techniques as forms of disturbance modeling. Integral control results from when seeks offset free control in the presence of a constant disturbance (c.f. Davison (1973*b*, 1973*a*)) and loop transfer recovery results when one models input disturbances. Disturbance modeling and offset free control have been discussed only briefly in the context of MPC (e.g. (Muske and Rawlings 1993, Rawlings, Meadows and Muske 1994)), and understanding their interaction appears to be promising direction of research. One proposal is to understand the (convex) geometry of linear MPC and use related concepts from (linear) geometric control theory (Wonham 1985).

Chapter 10

Conclusions

In this dissertation we investigated the moving horizon approximation as an online optimization strategy for the constrained process monitoring and control of nonlinear discrete-time systems. A framework was proposed for analyzing the stability properties of the moving horizon approximation. This framework allowed us to derive sufficient conditions for stability and propose practical algorithms for online implementation. This framework should prove useful for the future development of online optimization as a tool for constrained estimation and control. The main contribution of the dissertation and the basis of the proposed framework is the set of dual inequalities

$$\hat{\mathcal{Z}}_j(p) \leq \min_{x_{j-N}, \{w_k\}_{k=j-N}^{j-1}} \left\{ \sum_{k=j-N}^{j-1} L_k(w_k, v_k) + \hat{\mathcal{Z}}_{j-N}(x_{j-N}) \ : \ x_j = p \right\}$$

and

$$F_j(p) \geq \min_{\{u_k\}_{k=j}^{j+N-1}} \left\{ \sum_{k=j}^{j+N-1} l_k(u_k, x_k) + F_{j+N}(x_{j+N}) \ : \ x_j = p \right\}$$

discussed in Chapter 3. These dual inequalities constitute the core of stability results and also motivate practical algorithms for online implementation. Though this dissertation focused primarily on analyzing stability properties of the moving horizon approximation, practical and computational issues related to moving horizon estimation (MHE) and model predictive control (MPC) were also investigated. These results are briefly summarized below.

Chapter 2 investigated the issues regarding inequality constraints in process monitoring. One can significantly improve the quality of state estimates for certain problems by incorporating prior knowledge in the form of inequality constraints. Inequality constraints provide a flexible tool for complementing process knowledge and as a strategy also for model simplification. Chapter 3 developed a general theory for MHE and proposed algorithms for online implementation. Chapter 4 applied these results to the case when the system is linear, the objectives are quadratic, and the constraints are polyhedral convex sets.

Chapter 5 reviewed the basic theory of nonlinear MPC and discussed algorithms for online implementation. Chapter 6 discussed linear MPC and established techniques for handling inequality constraints active at steady state, a case that has not been treated in previous MPC theory. Through a series of examples, we showed how this case is significant in applications.

The practicality of MPC is partially limited by the ability to solve optimization problems in real time. This requirement limits the viability of MPC as a control strategy for large problems. One strategy for improving the computational performance is to formulate MPC using linear programming. Chapter 7 explored linear programming formations of MPC and demonstrated how the nonsmoothness of the objective function may yield either dead-beat or idle control response. Chapter 8 presented a structured interior-point method for the efficient solution of the optimal control problem in MPC. The cost of this approach is linear in the horizon length, compared with cubic growth for a naive approach. We also investigated strategies for further decomposing the problem structure in sheet and film forming processes.

Chapter 9 addressed the issues of output feedback and robustness by formulating MPC as a dynamic game. The game formulation allowed us to obtain a separation for output feedback and prove that the closed-loop system has finite gain. Furthermore, the added cost associated with formulating MPC as a dynamic game is negligible; the resulting problem is a quadratic program, though the optimization problem is no longer sparse. These results, however, are extremely conservative. Limitations of proposed strategy were discussed.

Suggestions for Future Work

There are still many unresolved issues in MHE and MPC. A brief list of some of these exciting research problems are described below.

- The strength and weakness of MHE and MPC is the use of online optimization. For most linear processes, the optimization problems can be reliably solved in less than 1 second on desktop computers using standard software. However, for nonlinear processes, computational difficulties often arise when one attempts to solve the optimization problems online. Significant progress has been made in developing efficient and robust algorithms for generic optimal control problems (e.g. (Biegler 1997)). However, an open question is how to incorporate these algorithms with the suboptimal MHE and MPC strategies described in Chapters 3 and 5, where the objective is to find a feasible solution rather a (locally) optimal solution. Only after a feasible solution is found and time permits may the optimization algorithm proceed to search for a (local) solution. One proposal is to tailor the algorithm described by Panier and Tits (1993) to the optimal control problems arising in MHE and MPC.

- Monitoring

  Most research in control is directed towards designing controllers with desired performance characteristics. Once the controller has been commissioned, the immediate question is whether the controller achieves and consistently maintains the desired objectives. It is difficult for the engineer to reliably monitor and diagnose the controller directly from raw data (Kozub 1997). Consequently, tools are needed to determine whether a controller satisfies its performance objectives and, if necessary, diagnose the problem when it does not. A handful of tools exist for monitoring univariate controllers using minimum variance estimates as a benchmark (Åström 1970, Harris 1989, Desborough and Harris 1992, Qin 1998). However, these tools become unwieldy when extended to multivariable systems. Furthermore, it is not clear whether these methods can be extended to MPC, which is a nonlinear controller. *Simple* tools, therefore, are needed to effectively monitor and diagnose multivariable control systems, in particular model predictive control. Monitoring is still a nascent field, and it remains to be seen how these problems will be satisfactorily resolved.

- Robustness

  Designing controllers robust to model uncertainty is still an open research problem. Standard methods based on game theory, or worst cast analysis, do not appear amenable to MPC as illustrated in Chapter 9. Alternative techniques are required to quantify robustness in MPC. However, it is not clear how to solve this problem systematically without resorting to dynamic programming.

In our opinion a more promising avenue of research is to devise a set of examples, pathological if necessary, for which MPC fails. By diagnosing specific examples, one can understand where and why robustness problems arise in MPC and then modify the algorithm accordingly. We believe this course of action is far more likely to yield worthwhile proposals than trying to obtain a general solution. Currently no such examples appear to exist in the literature.

- Output Feedback and Stability

  A controller consists of two parts: the estimator and the regulator. One typically establishes stability of the estimator and regulator separately and then tacitly assumes the combined system is stable. In Chapter 9 we demonstrated using game theory how the one can prove linear MPC is stable with output feedback. One can establish similar results also using probability theory. However, extending these results to nonlinear systems is difficult as one needs to resolve many technical details. An alternative and potentially simpler approach is to prove that the regulator is stable with decaying perturbations (Luenberger 1966, Vidyasagar 1980, Scokaert, Rawlings and Meadows 1997). These arguments require a converse Lyapunov theorem. To satisfy the conditions of a converse Lyapunov theorem, the regulator needs to be Lipschitz continuous. However, it is well known that MPC may yield discontinuous feedback (Meadows et al. 1995). One, therefore, needs to develop conditions for which the regulator is Lipschitz continuous or relax the Lipschitz condition.

- Adaptive Control

  Adaptive control is the "Holy Grail" of control theory. This is and will remain an exciting research problem as it implicitly involves all of the issues discussed above.

Appendix A

Existence

Standard existence arguments for finite-horizon optimal control problems typically employ the Weierstrass Maximum Theorem in conjunction with coercivity arguments. These arguments rarely extend to infinite-horizon optimal control problems. Instead, one needs to employ alternative arguments. Keerthi and Gilbert (1985) provide existence results for infinite-horizon optimal control problems under minimal assumptions, though the arguments are quite complicated. One may also establish existence by analyzing the problem in a weak* topology, thereby relaxing the restrictive compactness assumption with Alaoglu's Theorem (c.f. Luenberger (1969)). The purpose of this Appendix is to provide a "simple" proof for existence of an infinite-horizon, constrained, linear quadratic control problem. In particular we are concerned with conditions for the existence of a solution to the problem

$$\Phi^* = \min_{\{u_k, x_k\}_{k=0}^{\infty}} \Phi(\{u_k, x_k\}) \tag{A.1}$$

subject to the constraints

$$x_0 = \bar{x}, \qquad x_{k+1} = Ax_k + Bu_k, \tag{A.2a}$$

$$u_k \in \mathbb{U}, \qquad x_k \in \mathbb{X}, \tag{A.2b}$$

and the objective function

$$\Phi(\{u_k, x_k\}) = \sum_{k=0}^{\infty} x_k^T Q x_k + u_k^T R u_k,$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, and the sets $\mathbb{U}$ and $\mathbb{X}$ are closed and convex.

Definition A.0.1 *We say the sequence* $\{\bar{u}_k, \bar{x}_k\}_{k=0}^{\infty}$ *is* **admissible** *if the sequence* $\{\bar{u}_k, \bar{x}_k\}_{k=0}^{\infty}$ *satisfies the constraints (A.2) and* $\Phi(\{u_k, x_k\}) < \infty$.

Theorem A.0.2 *Suppose the matrices Q and R are positive definite. If an admissible solution* $\{\bar{u}_k, \bar{x}_k\}_{k=0}^{\infty}$ *exists, then there exists a unique solution to (A.1).*

Before proving Theorem A.0.2, we first prove the following Lemma.

Lemma A.0.3 *Suppose the matrix Q is positive definite, the set $\mathbb{Z}$ is a closed and convex, and*

$$\Theta(z_1) = \min_z \{\Theta(z) \;:\; z \in \mathbb{Z}\},$$

where $\Theta(z) = z^T Q z$. *Let* $z_2 = z_1 + \Delta \in \mathbb{Z}$. *If* $\Theta(z_2) - \Theta(z_1) \le \epsilon$ *for* $\epsilon > 0$, *then*

$$\|\Delta\| \le \sqrt{\frac{\epsilon}{\lambda_{\min}(Q)}}.$$

Proof. We know

$$\Theta(z_2) - \Theta(z_1) = 2z_1^T Q\Delta + \Delta^T Q\Delta.$$

By optimality, $\nabla\Theta(z)^T\Delta \geq 0$. Otherwise, if we consider $\Delta' = \gamma\Delta$ for $\delta << 1$, then we have

$$\Theta(z_1 + \Delta') - \Theta(z_1) = \underbrace{\gamma\nabla\Theta(z_1)^T\Delta}_{<0,\ O(\|\gamma\|)} + \underbrace{\gamma^2\Delta^T Q\Delta}_{\geq 0,\ O(\|\gamma\|^2)}.$$

Since $\mathbb{Z}$ is convex, there exists a feasible descent direction for γ small, thereby contradicting optimality.

So, we obtain the following inequality

$$\epsilon \geq \Theta(z_1 + \Delta) - \Theta(z_1) \geq \Delta^T Q\Delta \geq \lambda_{\min}\|\Delta\|^2.$$

This inequality implies

$$\sqrt{\frac{\epsilon}{\lambda_{\min}(Q)}} \geq \|\Delta\|,$$

and the lemma follows as claimed. □

Proof. [Theorem A.0.2] Let

$$\Phi_\infty = \sum_{k=0}^{\infty} \bar{x}_k^T Q\bar{x}_k + \bar{u}_k^T R\bar{u}_k.$$

We know by assumption that $\Phi_\infty < \infty$, because the sequence $\{\bar{u}_k, \bar{x}_k\}_{k=0}^{\infty}$ is admissible. Consider the finite horizon problem

$$\Phi_N^* = \min_{\{u_k, w_k\}_{k=0}^{N}} \sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k, \tag{A.3}$$

subject to the constraints (A.2). A solution exists to the problem (A.3) because the sequence $\{\bar{u}_k, \bar{x}_k\}_{k=0}^{N}$ is admissible for the finite-horizon problem and the cost functions are quadratic (Frank and Wolfe 1956). Furthermore, the solution is unique, because the objective function is strictly convex. Let $\{u_{k|N}, x_{k|N}\}_{k=0}^{N}$ denote the solution to (A.3). Optimality implies

$$\Phi_N^* \leq \Phi_{N+1}^* \leq \Phi_\infty.$$

Because sequence $\{\Phi_N^*\}$ is monotone nondecreasing and bounded above by Φ_∞, the sequence converges to some real number Φ_∞^*.

Let

$$U = \inf_{\{u_k, x_k\}_{k=0}^{\infty}} \Phi(\{u_k, x_k\})$$

subject to the constraints (A.2). Note that $M \geq \Phi_N^*$. Suppose $\Phi_\infty^* > U$. This supposition implies that there exists an $\epsilon > 0$ such that $\Phi_\infty^* - U \geq \epsilon$. Choose $\bar{N}$ such that $\Phi_\infty^* - \Phi_N^* < \epsilon$ for all $N \geq \bar{N}$. We have, therefore, the following contradiction for all $N \geq \bar{N}$:

$$\epsilon > \Phi_\infty^* - \Phi_N^* = \Phi_\infty^* - U + U - \Phi_N^* \geq \Phi_\infty^* - U \geq \epsilon.$$

Hence, $\Phi_\infty^* = U$ and $\Phi_\infty^* = \Phi^*$.

Consider the Banach space $l_2^\infty(\mathbb{R}^m \times \mathbb{R}^n)$. Let, for $M > N$,

$$\{u_{k|N}, x_{k|N}\}_{k=0}^{M} := \{u_{k|N}, x_{k|N}\}_{k=0}^{N} \times \{0^m, 0^n\}_{k=N+1}^{M}.$$

Let $\delta > 0$ and choose $\epsilon > 0$ such that

$$\sqrt{\frac{\epsilon}{[\lambda_{\min}(R) \wedge \lambda_{\min}(Q)]}} \leq \frac{\delta}{2}.$$

Choose $\bar{N} > 0$ such that, for all $M > N \geq \bar{N}$,

$$\Phi_M^* - \Phi_N^* \leq \epsilon.$$

It follows immediately that

$$0 \leq \sum_{k=0}^{N} x_{k|M}^T Q x_{k|M} + u_{k|M}^T R u_{k|M} - \Phi_N^* \leq \epsilon \tag{A.4}$$

and

$$\sum_{k=N+1}^{M} x_{k|M}^T Q x_{k|M} + u_{k|M}^T R u_{k|M} \leq \epsilon.$$

By Lemma A.0.3, (A.4) implies

$$\|\{u_{k|M}, x_{k|M}\}_{k=0}^{N} - \{u_{k|N}, x_{k|N}\}_{k=0}^{N}\| \leq \frac{\delta}{2}.$$

Likewise, we have

$$\begin{aligned}
\epsilon \quad &\geq \quad \sum_{k=N+1}^{M} x_{k|M}^T Q x_{k|M} + u_{k|M}^T R u_{k|M}, \\
&\geq \quad \sum_{k=N+1}^{M} \lambda_{\min}(Q)\|x_{k|M}\|^2 + \lambda_{\min}(R)\|u_{k|M}\|^2, \\
&\geq \quad [\lambda_{\min}(R) \wedge \lambda_{\min}(Q)]\, \|\{x_{k|M}, u_{k|M}\}_{k=N+1}^{M}\|^2.
\end{aligned}$$

This inequality implies

$$\frac{\delta}{2} \geq \|\{x_{k|M}, u_{k|M}\}_{k=N+1}^{M}\|.$$

Combining the results, we have

$$\begin{aligned}
&\|\{x_{k|M}, u_{k|M}\}_{k=0}^{M} - \{x_{k|N}, u_{k|N}\}_{k=0}^{N}\| \leq \\
&\quad \|\{x_{k|M}, u_{k|M}\}_{k=0}^{N} - \{x_{k|N}, u_{k|N}\}_{k=0}^{N}\| + \|\{x_{k|M}, u_{k|M}\}_{k=N+1}^{M}\|, \\
&\quad \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.
\end{aligned}$$

So, we have a Cauchy sequence in a closed subset (the feasible region) of a Banach space. Hence, there is a unique limit point $\{x_{k|\infty}, u_{k|\infty}\}_{k=0}^{\infty}$ contained within the feasible region.

We know

$$\Phi^* = \sum_{k=0}^{\infty} x_{k|\infty}^T Q x_{k|\infty} + u_{k|\infty}^T R u_{k|\infty},$$

because the cost function is continuous. Otherwise, we would contradict $\Phi_N^* \to \Phi^*$. Hence, a unique solution exists as claimed. $\square$

Appendix B

Duality

In this appendix, we discuss the duality between unconstrained linear quadratic control and least squares estimation (i.e the Kalman filter). In Section 3.6, we established the dual stability condition for receding horizon control (RHC) and moving horizon estimation (MHE). Here, we establish the duality between the cost functions in the classic sense. The results established here are well known, though rarely are they established explicitly. We, therefore, establish the dual relationship between linear control and estimation using some elementary results from convex programming theory.

B.1 Some Preliminaries

Duality is a rich field that arises in many different areas. An excellent survey of duality is given by Luenberger (1992). In this appendix we focus on duality in convex (quadratic) programming. The interested reader is referred to Luenberger (1969), Rockafellar (1970), and Mangasarian (1994) for a further discussion of duality.

Given two Hilbert spaces X and Y, we say the mapping $A : X \to Y$ is **linear** if $A(x+y) = Ax+Ay$ for all $x, y \in X$ and $A(\alpha x) = \alpha Ax$ for all $\alpha \in \mathbb{R}$ and $x \in X$ and **bounded** if there exists $M \geq 0$ such that $\|Ax\| \leq M\|x\|$. We denote the set of all bounded linear mappings from X to Y as $bl(X, Y)$. We say f is a **bounded linear functional** if it is a bounded linear map from X to $\mathbb{R}$. The set of all bounded linear functionals is itself a vector space. This space is called the **dual** of X and is denoted by X^*. If $X = \mathbb{R}^n$, then it is straightforward to show $X^* = \mathbb{R}^n$ and $f^*(x) = \langle f^*, x\rangle$ where $f^* \in \mathbb{R}^n$.

Likewise, we can consider bounded linear functionals on the space X^*. The set of all bounded linear functionals on X^* is called the second dual and denoted by X^{**}. As $x \in X$ defines a bounded linear functional on X^* through the definition $x(f) := f(x)$, we have $X \subset X^{**}$. If X is a Hilbert space, then $X = X^{**}$ (i.e. Hilbert spaces are reflexive). Let X and Y be Hilbert spaces and $A \in bl(X, Y)$. The adjoint operator $A^* : Y^* \to X^*$ is defined by the equation

$$\langle x, A^*y^*\rangle = \langle Ax, y^*\rangle.$$

Note $I^* = I$ and $A^{**} = A$. When $A \in \mathbb{R}^{n\times m}$, then $A^* = A^T$.

Let X and Y be Hilbert spaces. Consider the following convex program

$$\inf_{x\in\mathbb{X}} \{f(x) \ : \ \Omega(x) = 0\},$$

where set $\mathbb{X} \subset X$ is convex, the functional $f : \mathbb{X} \to \mathbb{R}$ is convex, and the mapping $\Omega : \mathbb{X} \to Y$ is affine. Now consider the **primal functional**

$$\omega(z) = \inf_{x\in\mathbb{X}} \{f(x) \ : \ \Omega(x) = z\}.$$

Associated with the primal function defined on Y is the **dual functional**

$$\varphi(z^*) = \inf_{x \in \mathbb{X}} \{f(x) + \langle G(x), z^* \rangle\}$$

defined on Y^*. The following theorem relates the primal and the dual functional.

Theorem B.1.1 *Suppose the functional $f(\cdot)$ is convex, the mapping $G(\cdot)$ is affine, and the set $\mathbb{X}$ is convex. If*

$$\mu = \inf_{x \in \mathbb{X}} \{f(x) \ : \ \Omega(x) = 0\}$$

is finite and there exists no p such that $\langle p, G(\cdot) \rangle = 0$, then

$$\inf_{x \in \mathbb{X}} \{f(x) \ : \ \Omega(x) = 0\} = \max_{z^*} \varphi(z^*),$$

and the maximum is achieved for some z_0^*. If the infimum on the left is achieved by some $x_0 \in \mathbb{X}$, then

$$\langle \Omega(x_0), z_0^* \rangle = 0,$$

and x_0 minimizes $f(x) + \langle \Omega(x), z_0^* \rangle$ for some $x \in \mathbb{X}$.

Proof. The theorem is a special case of result established in Luenberger (1969) (see Theorem 1, p225) where the generalized Slater constraint qualification has been replaced with the generalized Karlin constraint qualification (Mangasarian 1994). □

B.2 The Dual System

The triple (A, B, C) defines a (finite-dimensional) linear mapping G from $\bar{u} = (x_0, u_0, \dots, u_{N-1})$ to $\bar{y} = (y_0, \dots .y_{N-1}, x_N)$ through the equations

$$x_{k+1} = Ax_k + Bu_k, \qquad y_k = Cx_k.$$

The adjoint (dual system) G^* maps from $\bar{u}^* = (u_1^*, \dots, u_N^*, x_N^*)$ to $(x_0^*, y_1^*, \dots, y_N^*)$, where G^* is defined by the equations

$$x_k^* = A^T x_{k+1}^* + C^T u_k^*, \qquad y_k^* = B^T x_k^*.$$

To prove the triple (A^T, C^T, B^T) defines the dual system, we use the definition of the adjoint mapping:

$$\langle Gu, u^* \rangle = \langle u, G^* u^* \rangle.$$

Substituting in, we obtain the following equality that establishes duality.

$$\begin{aligned}
\langle G\bar{u}, \bar{u}^* \rangle &= \langle x_N, x_N^* \rangle + \sum_{k=0}^{N-1} \langle y_k, u_{k+1}^* \rangle \\
&= \langle Ax_{N-1} + Bu_{N-1}, x_N^* \rangle + \sum_{k=0}^{N-1} \langle Cx_k, u_{k+1}^* \rangle \\
&= \langle x_{N-1}, A^T x_N^* + C^* u_N^* \rangle + \langle u_{N-1}, B^T x_N^* \rangle + \sum_{k=0}^{N-2} \langle x_k, C^T u_{k+1}^* \rangle, \\
&= \langle u_{N-1}, y_N^* \rangle + \langle x_{N-1}, x_{N-1}^* \rangle + \sum_{k=0}^{N-2} \langle x_k, C^T u_{k+1}^* \rangle, \\
&\;\;\vdots \\
&= \langle x_0, x_0^* \rangle + \sum_{k=0}^{N-1} \langle u_k^*, y_{k+1}^* \rangle, \\
&= \langle \bar{u}, G^* \bar{u}^* \rangle.
\end{aligned}$$

For a further discussion of duality in linear systems, see Callier and Desoer (1991).

B.3 Main Result

We first define the primal control problem and then establish that the estimation problem is the dual. For simplicity, and without loss of generality, we assume that the target in the control problem is zero and that the data are zero (i.e. $\{y_k\} = 0$). The triple (A, B, C) defines the following finite-horizon control problem

$$V(\bar{x}) = \min_{\pi} \{\phi(\pi) \;:\; \Omega(\pi) = 0, \;\; x_0 = \bar{x}\},$$

where $\pi_N = \left(\{x_k, e_k\}_{k=0}^N, \{u_k\}_{k=0}^{N-1}\right)$, the objective function is defined as

$$\phi(\pi) = \frac{1}{2} \sum_{k=0}^{N-1} e_k^T Q e_k + u_k^T R u_k + \frac{1}{2} e_N^T P e_N,$$

and the constraint function $\Omega(\cdot)$ is defined as

$$\Omega(\pi) = \begin{bmatrix} Ax_0 + Bu_0 - x_1 \\ Ax_1 + Bu_1 - x_2 \\ \vdots \\ Ax_{N-1} + Bu_{N-1} - x_N \\ Cx_0 + e_0 \\ Cx_1 + e_1 \\ \vdots \\ Cx_{N-1} + e_{N-1} \\ x_N + e_N \end{bmatrix}.$$

It can be shown that the constraint function satisfies the constraint qualification; i.e. the rows of $\Omega(\cdot)$ are linearly independent. If we assume $X = \mathbb{R}^{(N+1)n \times Nm}$ and

$$\mathbb{X} = \{\pi \;:\; x_0 = \bar{x}\},$$

then the primal functional is given by

$$\omega(z) = \min_{\pi \in \mathbb{X}} \{\phi(\pi) \ : \ \Omega(\pi) = z\}.$$

As $\mathbb{X}$ is of finite dimension and the function $\phi(\cdot)$ is quadratic, we can replace the "inf" with a "min." One may view the primal problem as the "cost to go" in optimal control. Likewise, as we show, one may view the dual problem as the "arrival cost" in estimation. We define the dual functional as

$$\varphi(z^*) = \min_{\pi \in \mathbb{X}} \{\phi(\pi) + \langle \Omega(\pi), z^* \rangle\}.$$

We now derive an analytic expression for the dual functional $\varphi(\cdot)$. Let $z^* = (\{\lambda_k\}_{k=1}^N, \{p_k\}_{k=0}^N)$ and

$$\begin{aligned}\mathcal{L}(\pi, z^*) &:= \phi(\pi) + \langle \Omega(\pi), z^* \rangle \\ &= \frac{1}{2}\sum_{k=0}^{N-1} e_k^T Q e_k + u_k^T R u_k + \frac{1}{2} e_N^T P e_N + \sum_{k=0}^{N-1} \lambda_{k+1}^T (Ax_k - Bu_k - x_{k+1}) + \lambda_0^T (\bar{x} - x_0) + \\ &\sum_{k=0}^{N-1} p_k^T (Cx_k + e_k) + p_N^T (x_N + e_N).\end{aligned}$$

Evaluating the partial derivatives, we obtain

$$\begin{aligned}\frac{\partial \mathcal{L}(\cdot)}{\partial u_k} &= Ru_k + B^T \lambda_{k+1}, \\ \frac{\partial \mathcal{L}(\cdot)}{\partial e_k} &= Qe_k + p_k, \\ \frac{\partial \mathcal{L}(\cdot)}{\partial e_N} &= Pe_N + p_N, \\ \frac{\partial \mathcal{L}(\cdot)}{\partial x_k} &= A^T \lambda_{k+1} - \lambda_k + C^T p_k, \\ \frac{\partial \mathcal{L}(\cdot)}{\partial x_N} &= -\lambda_N + p_N.\end{aligned}$$

If we set the partial derivatives equal to zero, then we obtain the equalities

$$u_k = -R^{-1} B^T \lambda_{k+1}, \qquad e_k = -Q^{-1} p_k, \qquad e_N = -P^{-1} p_N.$$

Substituting in for u_k and e_k, we obtain the following expression for $\mathcal{L}(\cdot)$:

$$\begin{aligned}\mathcal{L}(\pi, z^*) &= \frac{1}{2}\sum_{k=0}^{N-1} \lambda_{k+1} B R^{-1} B^T \lambda_{k+1} + p_k^T Q^{-1} p_k + \frac{1}{2} p_N^T P^{-1} p_N + \\ &\quad \sum_{k=0}^{N-1} \lambda_{k+1}^T (Ax_k + Bu_k - x_{k+1}) + \lambda_0^T (\bar{x} - x_0) + \sum_{k=0}^{N-1} p_k^T (Cx_k + e_k) + p_N^T (x_N + e_N), \\ &= \frac{1}{2}\sum_{k=0}^{N-1} \lambda_{k+1} B R^{-1} B^T \lambda_{k+1} + p_k^T Q^{-1} p_k + \frac{1}{2} p_N^T P^{-1} p_N + \\ &\quad \sum_{k=0}^{N-1} \lambda_{k+1}^T B u_k + x_k^T (A^T \lambda_{k+1} + C^T p_k - \lambda_k) + x_N^T (p_N - \lambda_N) + \sum_{k=0}^{N} p_k^T e_k + \lambda_0^T \bar{x}, \\ &= -\frac{1}{2}\sum_{k=0}^{N-1} \left(\lambda_{k+1}^T B R^{-1} B^T \lambda_{k+1} + p_k^T Q^{-1} p_k\right) - \frac{1}{2} p_N^T P^{-1} p_N + \lambda_0^T \bar{x},\end{aligned}$$

subject to the constraints

$$\lambda_N = p_N, \qquad \lambda_k = A^T\lambda_{k+1} + C^T p_k.$$

As λ_N and $\{p_k\}$ unique determine λ_k, we can represent the "adjoint" system as

$$\Omega^*(z^*) = 0,$$

where

$$\Omega^*(z^*) := \begin{bmatrix} A^T\lambda_N + C^T p_{N-1} - \lambda_{N-1} \\ \vdots \\ A^T\lambda_1 + C^T p_0 - \lambda_0 \\ B^T\lambda_N + v_N \\ B^T\lambda_{N-1} + v_{N-1} \\ \vdots \\ B^T\lambda_1 + v_1 \end{bmatrix}.$$

If we redefine $z = \left(\{p_k\}_{k=0}^N, \{\lambda_k, v_k\}_{k=1}^N\right)$, then we have

$$\varphi(z^*) = \left\{-\Phi^*(z^*) + \lambda_0^T\bar{x} \;:\; \Omega^*(z^*) = 0\right\},$$

where

$$\Phi^*(z^*) = \frac{1}{2}\sum_{k=0}^{N-1}\left(v_{k+1}^T R^{-1} v_{k+1} + p_k^T Q^{-1} p_k\right) + \frac{1}{2}p_N^T P^{-1} p_N.$$

Associated with the dual functional is the dual optimization problem (see Theorem B)

$$\max_{z^*} \varphi^*(z^*).$$

If we make the substitutions $j = N - k$, $x_j^* = \lambda_{N-k}$, and $u_j^* = p_{N-1-k}$, then

$$\max_{z^*} \varphi^*(z^*) = -\min_{z^*}\left(\frac{1}{2}\sum_{j=0}^{N-1} y_j^{*T} R^{-1} y_j^* + u_j^{*T} Q^{-1} u_j + \frac{1}{2}x_0^{*T} P^{-1} x_0^* - x_N^{*T}\bar{x}\right),$$

subject to the dual (adjoint) system

$$x_{j+1}^* = A^T x_j^* + C^T u_j^*, \qquad y_j^* = B^T x_j^*.$$

From the derivation of the Kalman filter (see Appendix D), we can recast the dual problem as

$$\min_{x_N^*} \frac{1}{2} x_N^* P_N^{-1} x_N^* - x_N^{*T}\bar{x},$$

where

$$P_{k+1} = CQC^T + A^T\left(P_k - P_k B(R + B^T P_k B)^{-1} B^T P_k\right) A,$$

subject to the initial condition $P_0 = P$. Solving the minimization analytically, we obtain

$$x_N^* = P_N\bar{x}.$$

Hence, the dual problem is equivalent to

$$\max_{z^*} \varphi(z^*) = \min_{z^*} \{\Phi^*(z^*) \ : \ \Omega^*(z^*) = 0, x_N^* = P_N \bar{x}\}.$$

Recall the triple (A, B, C) defines the estimation problem

$$Z(\bar{x}) = \min_{\mu} \{\theta(\mu) \ : \ \Sigma(\mu) = 0, \ \ x_N = \bar{x}\},$$

where $\mu = \left(\{x_k\}_{k=0}^N, \{w_k, v_k\}_{k=0}^{N-1}\right)$, the objective is defined as

$$\theta(\mu) := \sum_{k=0}^{N} w_k^T Q^{-1} w_k + v_k^T R^{-1} v_k + x_0^T \Pi^{-1} x_0,$$

and the constraint function $\Sigma(\cdot)$ is defined as

$$\Sigma(\mu) := \begin{bmatrix} Ax_0 - Bw_0 - x_1 \\ Ax_1 - Bw_1 - x_2 \\ \vdots \\ Ax_{N-1} + Bw_{N-1} - x_N \\ Cx_0 + v_0 \\ Cx_1 + v_1 \\ \vdots \\ Cx_N + v_N \end{bmatrix}$$

The dual problem, consequently, defines the adjoint estimation problem. The structure of the dual problem is equivalent in structure to the estimation problem except that the triple (A^T, C^T, B^T) now defines the problem. In particular, the "cost to go" in the control problem is equal to the adjoint "arrival cost" in estimation. This is our duality,

Let $Z^*(\cdot)$ denote the arrival cost for the adjoint system (A^T, C^T, B^T), then

$$\begin{aligned} V(\bar{x}) &= Z^*(P_N \bar{x}), \\ &= \bar{x}^T P_N P_N^{-1} P_N \bar{x} = \bar{x}^T P_N \bar{x}. \end{aligned}$$

We can also establish that the estimation problem is the dual of the control problem. If we let $V^*(\cdot)$ denote the cost to go for the adjoint system (A^T, C^T, B^T), then

$$\begin{aligned} Z(\bar{x}) &= V^*(\Pi_N^{-1} \bar{x}) \\ &= \bar{x}^T \Pi_N^{-1} \Pi_N \Pi_N^{-1} \bar{x} = \bar{x}^T \Pi_N^{-1} \bar{x}, \end{aligned}$$

where

$$\Pi_{k+1} = BQB^T + A\left(\Pi_k - \Pi_k C^T (R + C^T \Pi_k C^T)^{-1} C \Pi_k\right) A^T,$$

subject to the initial condition $\Pi_0 = \Pi$. Hence, the arrival cost is equal to the adjoint cost to go.

B.4 The Riccati Equation

The duality between control and estimation is well known from the properties of the associated Riccati equations. For control, the cost to go is given by

$$V(\bar{x}) = \bar{x}^T P_N \bar{x},$$

where the triple (A, B, C) defines the Riccati equation

$$P_{k+1} = C^T QC + A^T \left(P_k - P_k B^T (R + B^T P_k B) B P_k\right) A,$$

subject to the initial condition $P_0 = P$. Likewise, the arrival cost in estimation is given by

$$V(\bar{x}) = \bar{x}^T \Pi_N \bar{x},$$

where the triple (A, B, C) defines the Riccati equation

$$\Pi_{k+1} = BQB^T + A \left(\Pi_l - PC(R + C\Pi_k C^T) C^T \Pi_k\right) A^T,$$

subject to the initial condition $P_0 = \Pi$. One obtains, by inspection, the control Riccati equation from estimation Riccati equation if one uses the triple (A^T, C^T, B^T). Likewise, the estimation Riccati equation is equal to control Riccati equation if one uses the triple (A^T, C^T, B^T). Hence, the estimation and control Riccati equations are dual to each other.

Appendix C

A derivation of $\mathcal{H}_\infty$ control

C.1 Derivation of the $\mathcal{H}_\infty$ regulator

In this section we derive the solution to the dynamic game

$$V_N^0(x) = \min_{\{u_k\}_{k=0}^{N-1}} \max_{\{w_k\}_{k=0}^{N-1}} V_N(x_0, \{u_k\}, \{w_k\})$$

where

$$V_N(x_0, \{u_k\}, \{w_k\}) := \sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k + x_N^T P x_N - \gamma^2 w_k^T w_k$$

subject to the constraints

$$\begin{aligned} x_0 &= x \\ x_{k+1} &= Ax_k + Bu_k + Gw_k. \end{aligned}$$

First consider the following dynamic game

$$V_1(x_0) = \min_{x_0} \max_{x_0} \; V_1(x_0, u_0, w_0)$$

subject to the state equation

$$x_1 = Ax_0 + Bu_0 + Gw_0.$$

Substituting the state equation into the term $x_1^T P x_1$, we obtain the following expression

$$\begin{aligned} x_1^T P x_1 &= (Ax_0 + Bu_0 + Gw_0)^T P (Ax_0 + Bu_0 + Gw_0), \\ &= x_0^T A^T P A x_0 + u_0^T B^T P B u_0 + w_0^T G^T P G w_0 + \\ &\quad 2w_0^T G^T P (Ax_0 + Bu_0) + 2u_0^T B^T P A x_0. \end{aligned}$$

Substituting into the cost function, we obtain the new expression in terms of x_0, u_0, and w_0:

$$\begin{aligned} V_1(x_0, u_0, w_0) &= x_0^T (Q + A^T P A) x_0 + u_0^T (R + B^T P B) u_0 + w_0^T (G^T P G - \gamma^2) w_0 + \\ &\quad 2w_0^T G^T P (Ax_0 + Bu_0) + 2u_0^T B^T P A x_0. \end{aligned}$$

Solving the maximization first, we have the following expression for the partial derivative of V_1 with respect to w_0.

$$\frac{\partial V_1(x_0, u_0, w_0)}{\partial w_0} = (G^T P G - \gamma^2) w_0 + G^T P (Ax_0 + Bu_0).$$

In order for the solution to be well-posed, γ needs to satisfy the matrix inequality

$$(\gamma^2 - G^T P G) > 0. \tag{C.1}$$

Otherwise, the cost function is not strictly concave in w_0. Assuming γ satisfies (C.1), we can solve for the worst-case disturbance:

$$\left(\frac{\partial V_1(x_0, u_0, w_0)}{\partial w_0} = 0\right) \Longrightarrow w_0^* = (\gamma^2 - G^T P G)^{-1} G^T P (Ax_0 + Bu_0).$$

Substituting in for w_0^*, we have

$$\begin{aligned}
V_1(x_0, u_0, w_0^*) &= x_0^T (Q + A^T P A) x_0 + u_0^T (R + B^T P B) u_0 + \\
&\quad (Ax_0 + Bu_0)^T P G (G^T P G - \gamma^2)^{-1} G^T P (Ax_0 + B_0) + \\
&\quad 2(Ax_0 + Bu_0)^T P G (\gamma^2 - G^T P G)^{-1} G^T P (Ax_0 + B_0) + \\
&\quad 2u_0^T B^T P A x_0, \\
&= x_0^T (Q + A^T P A) x_0 + u_0^T (R + B^T P B) u_0 + \\
&\quad (Ax_0 + Bu_0)^T P G (\gamma^2 - G^T P G)^{-1} G^T P (Ax_0 + B_0) + \\
&\quad 2u_0^T B^T P A x_0.
\end{aligned}$$

Rearranging the cost function yields

$$\begin{aligned}
V_1(x_0, u_0, w_0^*) &= x_0^T (Q + A^T P A) x_0 + u_0^T (R + B^T P B) u_0 + \\
&\quad u_0^T B^T P G (\gamma^2 - G^T P G)^{-1} G^T P B u_0 + x_0^T A^T P G (\gamma^2 - G^T P G)^{-1} G^T P A x_0 + \\
&\quad 2u_0^T B^T \left(P G (\gamma^2 - G^T P G)^{-1} G^T P + P\right) A x_0.
\end{aligned}$$

Collecting terms yields the following expression for the cost function:

$$\begin{aligned}
V(x_0) &= x_0^T \left(Q + A^T \left(P + P G (\gamma^2 - G^T P G)^{-1} G^T P) G^T P\right) A\right) x_0 + \\
&\quad u_0^T \left(R + B^T \left(P + P G (\gamma^2 - G^T P G)^{-1} G^T P\right) B\right) u_0 + \\
&\quad 2u_0^T B^T \left(P G (\gamma^2 - G^T P G)^{-1} G^T P + P\right) A x_0.
\end{aligned}$$

We have the following identity using the matrix inversion lemma:

$$P G (\gamma^2 - G^T P G)^{-1} G^T P + P = \left(P^{-1} - \frac{1}{\gamma^2} G G^T\right)^{-1}.$$

Taking the partial derivative of $V_1(x_0, u_0, w_0^*)$ with respect to u_0 yields

$$\begin{aligned}
\frac{\partial V_1(x_0, u_0, w_0^*)}{\partial u_0} &= \left(R + B^T \left(P + P G (\gamma^2 - G^T P G)^{-1} G^T P)^{-1} G^T P\right) B\right) u_0 \\
&\quad B^T \left(P G (\gamma^2 - G^T P G)^{-1} G^T P + P\right) A x_0,, \\
&= \left(R + B^T \left(P^{-1} - \frac{1}{\gamma^2} G G^T\right)^{-1} B\right) u_0 + \\
&\quad B^T \left(P^{-1} - \frac{1}{\gamma^2} G G^T\right)^{-1} A x_0.
\end{aligned}$$

Solving the minimization, we obtain the following expression the optimal input

$$\left(\frac{\partial V_1(x_0, u_0, w_0^*)}{\partial u_0} = 0\right) \Longrightarrow$$

$$u_0^* = -\left(R + B^T \left(P^{-1} - \frac{1}{\gamma^2} G G^T\right)^{-1} B\right)^{-1} B^T \left(P^{-1} - \frac{1}{\gamma^2} G G^T\right)^{-1} A x_0.$$

Using the matrix inversion lemma once again yields

$$
\begin{aligned}
&\left(R + B^T\left(P^{-1} - \frac{1}{\gamma^2}GG^T\right)^{-1} B\right)^{-1} B^T\left(P^{-1} - \frac{1}{\gamma^2}GG^T\right)^{-1} = \\
&\quad R^{-1}B^T\left((P^{-1} - \frac{1}{\gamma^2}GG^T) + B^TR^{-1}B\right)^{-1}, \\
&= R^{-1}B^T\left(\left(I + (B^TR^{-1}B - \frac{1}{\gamma^2}GG^T)P\right)P^{-1}\right)^{-1}, \\
&= R^{-1}B^TP\left(I + (B^TR^{-1}B - \frac{1}{\gamma^2}GG^T)P\right)^{-1}.
\end{aligned}
$$

Hence, we have

$$
u_0^* = -R^{-1}B^TP\left(I + (B^TR^{-1}B - \frac{1}{\gamma^2}GG^T)P\right)^{-1} Ax_0.
$$

Recall, we have the following expression for the worst-case disturbance:

$$
\begin{aligned}
w_0^* &= (\gamma^2 - G^TPG)^{-1}G^TP(Ax_0 + Bu_0) \\
&= (\gamma^2 - G^TPG)^{-1}G^TP \cdot \\
&\quad \left(I - BR^{-1}B^TP\left(I + (B^TR^{-1}B - \frac{1}{\gamma^2}GG^T)P\right)^{-1}\right) Ax_0.
\end{aligned}
$$

Using the matrix inversion lemma, we have the following identity:

$$
\begin{aligned}
(\gamma^2 - G^TPG)^{-1}G^TP &= \frac{1}{\gamma^2}G^T(P^{-1} - \frac{1}{\gamma^2}GG^T)^{-1}, \\
&= \frac{1}{\gamma^2}G^TP(I - \frac{1}{\gamma^2}GG^TP)^{-1}.
\end{aligned}
$$

Making the following definition

$$
D := (I - \frac{1}{\gamma^2}GG^TP),
$$

we obtain the following identities:

$$
\begin{aligned}
&I - BR^{-1}B^TP\left(B^TR^{-1}BP + I - \frac{1}{\gamma^2}GG^T)P\right)^{-1} \\
&= I - BR^{-1}B^TP\left(B^TR^{-1}BP + D\right)^{-1}, \\
&= I - BR^{-1}B^TPD^{-1}\left(B^TR^{-1}BPD^{-1} + I\right)^{-1}, \\
&= (I + B^TR^{-1}BPD^{-1})^{-1},
\end{aligned}
$$

where the last inequality follows from the matrix inversion lemma. We have also the following identity:

$$
\begin{aligned}
D^{-1}(I + B^TR^{-1}BPD^{-1})^{-1} &= (D + B^TR^{-1}BP)^{-1} \\
&= \left(I + (B^TR^{-1}B - \frac{1}{\gamma^2}GG^T)P.\right)^{-1}
\end{aligned}
$$

So, we have the following expression for the worst-case disturbance:

$$\begin{aligned} w_0^* &= \frac{1}{\gamma^2} G^T P D^{-1} \left(I - BR^{-1}B^T P \left(B^T R^{-1} B P + D \right)^{-1} \right) A x_0, \\ &= \frac{1}{\gamma^2} G^T P \left(I + (B^T R^{-1} B - \frac{1}{\gamma^2} G G^T) P \right)^{-1} A x_0. \end{aligned}$$

Substituting u_0^* in $V(x_0, u_0, w_0^*)$, we obtain

$$\begin{aligned} V_1(x_0) &= x_0^T \left(Q + A^T (P^{-1} - \frac{1}{\gamma^2} G G^T)^{-1} A \right) x_0 - \\ & x_0^T A^T (P^{-1} - \frac{1}{\gamma^2} G G^T)^{-1} B \left(R + B^T (P^{-1} - \frac{1}{\gamma^2} G G^T)^{-1} B \right)^{-1} \cdot \\ & B^T (P^{-1} - \frac{1}{\gamma^2} G G^T)^{-1} A x_0. \end{aligned}$$

Making the definition

$$W := (P^{-1} - \frac{1}{\gamma^2} G G^T),$$

yields the following expressions for the optimal cost:

$$\begin{aligned} V_1(x_0) &= x_0^T \left(Q + A^T \left(W^{-1} - W^{-1} B (R + B^T W^{-1} B)^{-1} B^T W^{-1} \right) A \right) x_0, \\ &= x_0^T \left(Q + A^T (D + BR^{-1} B^T)^{-1} A \right) x_0, \\ &= x_0^T \left(Q + A^T \left((P^{-1} - \frac{1}{\gamma^2} G G^T) + BR^{-1} B^T \right)^{-1} A \right) x_0, \\ &= x_0^T \left(Q + A^T P \left(I + (BR^{-1} B^T - \frac{1}{\gamma^2} G G^T) P \right)^{-1} A \right) x_0, \end{aligned}$$

where the last inequality follows from the matrix inversion lemma.

We can extend the above result using standard dynamic programming arguments. Let

$$P_{k+1} = Q + A^T P_k \left(I + (BR^{-1} B^T - \frac{1}{\gamma^2} G G^T) P_k \right)^{-1} A,$$

subject to the initial condition $P_0 = P$. If

$$\gamma^2 - G^T P_k G > 0$$

for all $k = 0, \dots, (N-1)$, then

$$V_N(x_0) = x_0^T P_N x_0.$$

and

$$\begin{aligned} u_0^* &= -R^{-1} B^T P_N \left(I + (B^T R^{-1} B - \frac{1}{\gamma^2} G G^T) P_N \right)^{-1} A x_0, \\ w_0^* &= \frac{1}{\gamma^2} G^T P_N \left(I + (B^T R^{-1} B - \frac{1}{\gamma^2} G G^T) P_N \right)^{-1} A x_0. \end{aligned}$$

C.2 Derivation of $\mathcal{H}_\infty$ Estimator

Consider the problem

$$\mathcal{Z}_N(z) \max_{x_0,\{w_k\}_{k=0}^{N-1}} Z_N(x_0,\{w_k\})$$

where

$$Z_N(x_0,\{w_k\}) := \sum_{k=0}^{N-1} x_k^T Q x_k - \gamma^2 \left(w_k^T w_k + v_k^T v_k + (x_0-\bar{x})^T \Pi^{-1}(x_0-\bar{x}) \right)$$

subject to the constraints

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Gw_k, \\ y_k &= Cx_k + Ev_k, \\ x_N &= z. \end{aligned}$$

where the matrix E is nonsingular.

First consider the problem

$$\max_{x_0,w_0} Z_1(x_0,w_0)$$

subject to the constraints

$$x_1 = Ax_k + Bu_k + Gw_k, \;\; y_k = Cx_k + Ew_k, \;\; x_1 = z.$$

This problem is equivalent to

$$\min_{x_0} \; v_k^T v_k + (x_0-\bar{x})^T \Pi^{-1}(x_0-\bar{x}) - \frac{1}{\gamma^2} x_0^T Q x_0$$

Substituting the model equation, we obtain

$$\min_{x_0} \; (y_0 - Cx_0)^T (EE^T)^{-1}(y - Cx_0) + (x_0-\bar{x})^T \Pi^{-1}(x_0-\bar{x}) - \frac{1}{\gamma^2} x_0^T Q x_0$$

Rearranging and collecting terms, we obtain the following equivalent minimization

$$\min_{x_0} \; x_0^T \left(C^T (EE^T)^{-1} C + \Pi^{-1} - \frac{1}{\gamma^2} Q \right) x_0 - x_0^T \Pi_0^{-1} \bar{x} - x_0^T C^T (EE^T)^{-1} y_0.$$

If γ satisfies the matrix inequality

$$\gamma^2 \Pi^{-1} - Q > 0,$$

then the problem admits the solution

$$\begin{aligned} \hat{x}_0 &= \left(C^T (EE^T)^{-1} C + \Pi^{-1} - \frac{1}{\gamma^2} Q \right)^{-1} \left(\Pi_0^{-1} \bar{x} + C^T (EE^T)^{-1} y_0 \right), \\ &= \left(C^T (EE^T)^{-1} C + \Pi^{-1} - \frac{1}{\gamma^2} Q \right)^{-1} \times \\ &\quad \left(\left(C^T (EE^T)^{-1} C + \Pi^{-1} - \frac{1}{\gamma^2} Q \right) \bar{x} + \frac{1}{\gamma^2} Q \bar{x} + C (EE^T)^{-1} (y_0 - C\bar{x}) \right), \\ &= \bar{x} + \left(C^T (EE^T)^{-1} C + \Pi^{-1} - \frac{1}{\gamma^2} Q \right)^{-1} \left(\frac{1}{\gamma^2} Q \bar{x} + C (EE^T)^{-1} (y_0 - C\bar{x}) \right). \end{aligned}$$

Now consider the cost for an arbitrary x:

$$
\begin{aligned}
Z_{1|0}(x) &= (y_0 - Cx)^T(EE^T)^{-1}(y_0 - Cx) + (x-\bar{x})^T\Pi^{-1}(x-\bar{x}) - \frac{1}{\gamma^2}x^TQx, \\
&= (y_0 - C(x+\hat{x}_0-\hat{x}_0))^T(EE^T)^{-1}(y_0 - C(x+\hat{x}_0-\hat{x}_0)) + \\
&\quad (x+\hat{x}_0-\hat{x}_0-\bar{x})^T\Pi(x+\hat{x}_0-\hat{x}_0-\bar{x}) - \\
&\quad \frac{1}{\gamma^2}(x-\hat{x}_0+\hat{x}_0)^TQ(x-\hat{x}_0+\hat{x}_0), \\
&= (y - C\hat{x}_0 - C\Delta x)^T(EE^T)^{-1}(y - C\hat{x}_0 - C\Delta x) + \\
&\quad (\Delta x + \hat{x}_0 - \bar{x})^T\Pi^{-1}(\Delta x + \hat{x}_0 - \bar{x}) + \frac{1}{\gamma^2}(\Delta x + \hat{x}_0)^TQ(\Delta x + \hat{x}_0), \\
&= (y_0 - C\hat{x}_0)^T(EE^T)^{-1}(y - C\hat{x}_0) + \Delta x^TC^T(EE^T)^{-1}C\Delta x - \\
&\quad 2\Delta xC^T(EE^T)^{-1}(y_0 - C\hat{x}_0) + \Delta x^T\Pi^{-1}\Delta x + (\hat{x}_0-\bar{x})^T\Pi^{-1}(\hat{x}_0-\bar{x}) + \\
&\quad 2\Delta x\Pi^{-1}(\hat{x}_0-\bar{x}) - \frac{1}{\gamma^2}\Delta x^TQ\Delta x - \frac{1}{\gamma^2}\hat{x}_0^TQ\hat{x}_0 - \frac{1}{\gamma^2}\Delta x^TQ\hat{x}_0.
\end{aligned}
$$

By optimality we know

$$\Pi_0^{-1}(\hat{x}_0 - \bar{x}) - C^TR^{-1}(y_0 - C\hat{x}_0) - \frac{1}{\gamma}Q\hat{x}_0 = 0.$$

Hence,

$$Z_{1|0}(x) = \gamma^2(x-\hat{x}_0)^T\left(C^T(EE^T)^{-1}C + \Pi^{-1} - \frac{1}{\gamma^2}Q\right)(x-\hat{x}_0) - \alpha,$$

where

$$\alpha = \hat{x}^TQ\hat{x}_0 - \gamma^2\left((y_0 - C\hat{x}_0)^T(EE^T)^{-1}(y - C\hat{x}_0) + (\hat{x}_0-\bar{x})^T\Pi^{-1}(\hat{x}_0-\bar{x})\right).$$

Using the optimality principle, we know

$$V_1(z) = \alpha - \max_{x,w_0}\ w_0^Tw_0 + (x-\hat{x}_0)^T\left(C^T(EE^T)^{-1}C + \Pi^{-1} - \frac{1}{\gamma^2}Q\right)^{-1}(x-\hat{x}_0).$$

Let

$$\Sigma := C^T(EE^T)^{-1}C + \Pi^{-1} - \frac{1}{\gamma^2}Q.$$

Constructing the Lagrangian, we have

$$\mathcal{L}_1 = w_0^Tw_0 + \gamma^2(x-\hat{x}_0)^T\Sigma^{-1}(x-\hat{x}_0) + \lambda_0^T(Ax + Gw_0 - z).$$

The optimality conditions are

$$
\begin{gathered}
w_0 + G^T\lambda_0 = 0, \\
\gamma^2\Sigma^{-1}(x-\hat{x}_0) + A^T\lambda_0 = 0.
\end{gathered}
$$

Eliminating λ_0 and using the model equation, we obtain

$$w_0 = G\left(A\Sigma^{-1}A^T + GG^T\right)^{-1}(z - A\hat{x}_0)$$

and

$$(x_0 - \hat{x}_0) = \Sigma^{-1}A^T\left(A\Sigma^{-1}A^T + GG^T\right)^{-1}(z - A\hat{x}_0).$$

Substituting in for x_0 and w_0, we obtain

$$V_1(z) = \alpha - \gamma^2(z - A\hat{x}_0)\left(A\Sigma^{-1}A^T + GG^T\right)^{-1}(z - A\hat{x}_0) + \alpha.$$

If we make the definition

$$\Pi_1 = A\left(C^T(EE^T)^{-1}C + \Pi^{-1} - \frac{1}{\gamma^2}Q\right)^{-1} A^T + GG^T,$$

then

$$V_1(z) = \alpha - \gamma^2(z - A\hat{x}_0)\Pi_1^{-1}(z - A\hat{x}_0).$$

To complete the derivation of the estimator, we need only to employ forward dynamic programming and the results we derived above. If γ satisfies the matrix inequality

$$\gamma^2\Pi_k^{-1} - Q > 0$$

for $k = 0, \ldots, (N-1)$, then

$$V_N(z) = \alpha_N - \gamma^2(z - \bar{x}_N)\Pi_N^{-1}(z - \bar{x}_N)$$

where α_k is a positive constant,

$$\Pi_{k+1} = A\left(C^T(EE^T)^{-1}C + \Pi_k^{-1} - \frac{1}{\gamma^2}Q\right)^{-1} A^T + GG^T,$$

and

$$\bar{x}_{k+1} = A\bar{x}_k + Bu_k + A\left(C^T(EE^T)^{-1}C + \Pi_k^{-1} - \frac{1}{\gamma^2}Q\right)^{-1}\left(\frac{1}{\gamma^2}Q\bar{x}_k + C(EE^T)^{-1}(y_k - C\bar{x}_k)\right),$$

subject to the initial conditions $\Pi_0 = \Pi$ and $\bar{x}_0 = \bar{x}$.

Appendix D

The Kalman Filter

In this appendix we derive the Kalman filter using dynamic programming and block factorization. The latter approach may be used to extend the results of Chapter 8 to moving horizon estimation.

We can formulate the Kalman filter as the solution to the following least squares estimation problem

$$\phi_T = \min_{x_0,\, \{w_k\}_{k=0}^{T-1}} \sum_{k=0}^{T-1} w_k^T Q^{-1} w_k + v_k^T R^{-1} v_k + (x_0 - \hat{x}_0)^T \Pi_0^{-1} (x_0 - \hat{x}_0)$$

subject to the model equation

$$\begin{aligned} x_{k+1} &= Ax_k + Gw_k, \\ y_k &= Cx_k + v_k. \end{aligned}$$

Dynamic Programming

We proceed inductively. Let $T = 1$ and consider the problem

$$\min_{x_0,\, w_0} w_0^T Q^{-1} w_0 + v_0^T R^{-1} v_0 + (x_0 - \hat{x}_0)^T \Pi_0^{-1} (x_0 - \hat{x}_0).$$

This problem is equivalent to the minimization

$$\min_{x_0} v_0^T R^{-1} v_0 + (x_0 - \hat{x}_0)^T \Pi_0^{-1} (x_0 - \hat{x}_0).$$

The optimality conditions for this problem are

$$\left(C^T R^{-1} C + \Pi_0^{-1}\right) x_0 = C^T R^{-1} y_0 + \Pi_0^{-1} \hat{x}_0,$$

and the optimal solution is

$$x_0^* = \left(C^T R^{-1} C + \Pi_0^{-1}\right)^{-1} \left(C^T R^{-1} y_0 + \Pi_0^{-1} \hat{x}_0\right).$$

Using the matrix inversion lemma, we have equivalent solution

$$x_0^* - \hat{x}_0 = \Pi_0 C^T (R + C\Pi C^T)^{-1} (y_0 - C\hat{x}_0).$$

So, we obtain the solution

$$\hat{x}_1 = A\hat{x}_0 + A\Pi_0 C^T (R + C\Pi C^T)^{-1} (y_0 - C\hat{x}_0).$$

Let us now consider the cost associated with an arbitrary x_0 and w_0:

$$\begin{aligned}\phi(x_0, w_0) &= w_0^T Q^{-1} w_0 + v_0^T R^{-1} v_0 + (x_0 - \hat{x}_0)^T \Pi^{-1} (x_0 - \hat{x}_0), \\ &= w_0^T Q^{-1} w_0 + (y_0 - C(x_0^* + \Delta))^T R^{-1} (y_0 - C(x_0^* + \Delta)) + (x_0^* + \Delta)^T \Pi_0^{-1} (x_0^* + \Delta),\end{aligned}$$

where $\Delta = x_0 - x_0^*$. We can rearrange the above expression as follows

$$\begin{aligned}\phi(w_0, x_0) = \;& w_0^T Q^{-1} w_0 + \Delta^T (C^T R^{-1} C + \Pi_0^{-1}) \Delta + \\ & \Delta^T \left(C^T R^{-1} y_0 + \Pi_0^{-1} \hat{x}_0 - \left(C^T R^{-1} C + \Pi_0^{-1} \right) x_0 \right) + \\ & (y_0 - C x^*)^T R^{-1} (y_0 - C x^*) + (x_0^* - \hat{x}_0)^T \Pi_0^{-1} (x_0^* - \hat{x}_0).\end{aligned}$$

Using the optimality conditions, we obtain

$$\phi(x_0, w_0) = w_0^T R^{-1} w_0 + \Delta^T (C^T R^{-1} C + \Pi_0^{-1}) \Delta + \phi_1.$$

Now consider the problem

$$\mathcal{Z}_1(z) = \min_{x_0,\, w_0} \left\{ w_0^T Q^{-1} w_0 + v_0^T R^{-1} v_0 + (x_0 - \hat{x}_0)^T \Pi_0^{-1} (x_0 - \hat{x}_0) \;:\; x_1 = z \right\}.$$

This problem is equivalent to

$$\mathcal{Z}_1(z) = \min_{x_0,\, w_0} \left\{ \phi_1(x_0, w_0) \;:\; x_1 = z \right\}.$$

The optimality conditions for this problem are

$$\begin{gathered} Q^{-1} w_0 + G^T \lambda = 0, \\ (C^T R^{-1} C + \Pi_0^{-1})(x_0 - x_0^*) + A^T \lambda = 0, \\ A x_0 + G w_0 = z. \end{gathered}$$

Solving for x_0, we obtain

$$\hat{x}_1 - z = A (C^T R^{-1} C + \Pi_0^{-1})^{-1} A^T \lambda - G w_0.$$

Solving for w_0, we obtain

$$\begin{aligned}\hat{x}_1 - z &= \left(A (C^T R^{-1} C + \Pi_0^{-1})^{-1} A^T + G Q G^T \right) \lambda, \\ &= \left(G Q G^T + A (\Pi_0 - \Pi_0 C^T (R + C \Pi_0 C^T)^{-1} C \Pi_0) A^T \right) \lambda.\end{aligned}$$

If we make the definition

$$\Pi_1 = G Q G^T + A (\Pi_0 - \Pi_0 C^T (R + C \Pi_0 C^T)^{-1} C \Pi_0) A^T,$$

then

$$\begin{gathered} Q^{-1} w_0 = G^T \Pi^{-1} (z - \hat{x}_1), \\ (C^T R^{-1} C + \Pi_0^{-1})(x_0 - x_0^*) = A^T \Pi_1^{-1} (z - \hat{x}_1). \end{gathered}$$

Substituting in these expressions, we obtain

$$\mathcal{Z}_1(z) = (z - \hat{x}_1)^T \Pi_1^{-1} (z - \hat{x}_1) + \phi_1.$$

Now consider arbitrary T and assume

$$\mathcal{Z}_{T-1}(z) = (z - \hat{x}_{T-1})^T \Pi_{T-1}^{-1} (z - \hat{x}_{T-1}) + \phi_{T-1},$$

where Π_k and $\hat{x}_k$ are given by the recursive expressions

$$\Pi_{k+1} = GQG^T + A(\Pi_k - \Pi_k C^T (R + C\Pi_k C^T)^{-1} C\Pi_k) A^T, \tag{D.1}$$

and

$$\hat{x}_{k+1} = A\hat{x}_k + A\Pi_k C^T (R + C\Pi_k C^T)^{-1} (y_k - C\hat{x}_k).$$

Using principle of optimality and the induction hypothesis, we know

$$\phi_T = \min_{x_{T-1},\, w_{T-1}} w_{T-1}^T Q^{-1} w_{T-1} + v_{T-1}^T R^{-1} v_{T-1} + (x_{T-1} - \hat{x}_{T-1})^T \Pi_{T-1}^{-1} (x_{T-1} - \hat{x}_{T-1}) + \phi_{T-1}.$$

Hence,

$$\hat{x}_T = A\hat{x}_{T-1} + A\Pi_T C^T (R + C\Pi_{T-1} C^T)^{-1} (y_{T-1} - C\hat{x}_{T-1})$$

and

$$\mathcal{Z}_T(z) = (z - \hat{x}_T)^T \Pi_T^{-1} (z - \hat{x}_T) + \phi_T,$$

where Π_T is given by (D.1).

Batch Least Squares

Consider $T = 1$ and the equivalent problem

$$\min_{x_0} v_0^T R^{-1} v_0 + (x_0 - \hat{x}_0)^T \Pi_0^{-1} (x_0 - \hat{x}_0).$$

The optimality conditions are

$$\begin{bmatrix} \Pi_0^{-1} & C^T & \\ C & & -I \\ & -I & R^{-1} \end{bmatrix} \begin{bmatrix} x_0 \\ \lambda \\ v_0 \end{bmatrix} = \begin{bmatrix} \Pi_0^{-1} \hat{x}_0 \\ y_0 \\ 0 \end{bmatrix}.$$

Block elimination gives:

$$x_0 = (\Pi_0^{-1} + C^T R^{-1} C)^{-1} (\Pi_0^{-1} \hat{x}_0 + C^T R^{-1} y_0).$$

From Matrix Inversion Lemma, we have the following equalities;

$$\begin{aligned} (\Pi_0^{-1} + C^T R^{-1} C)^{-1} &= \Pi_0 - \Pi_0 C^T (R + C\Pi_0 C^T)^{-1} C Q_0, \\ (\Pi_0^{-1} + C^T R^{-1} C)^{-1} C^T R^{-1} &= \Pi_0 C^T (C\Pi_0 C^T + R)^{-1}. \end{aligned}$$

This gives us the well known result for the Kalman filter:

$$x_0 = \hat{x}_0 + \Pi_0 C^T (C\Pi_0 C^T + R)^{-1} (y_0 - C\hat{x}_0).$$

Now consider the case for $T = 2$, which we use to generalize for all T. The solution of this problem is again obtained from the Karush-Kuhn-Tucker conditions:

$$\begin{bmatrix} \Pi_0^{-1} & C^T & & & A^T C^T & \\ C & & -I & & & \\ & -I & R^{-1} & & & \\ & & & Q^{-1} & G^T C^T & \\ CA & & & CG & & -I \\ & & & & -I & R^{-1} \end{bmatrix} \begin{bmatrix} x_0 \\ \lambda_0 \\ v_0 \\ w_0 \\ \lambda_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} \Pi_0^{-1}\hat{x}_0 \\ y_0 \\ 0 \\ 0 \\ y_1 \\ 0 \end{bmatrix}.$$

Block elimination of the first stage gives:

$$\begin{bmatrix} \Pi_0^{-1} & A^T C^T & \\ CA & -CGQG^T C^T & -I \\ & -I & R^{-1} \end{bmatrix} \begin{bmatrix} x_0 \\ \lambda_1 \\ x_1 \end{bmatrix} = \begin{bmatrix} \Pi_0^{-1}\hat{x}_0 \\ y_1 \\ 0 \end{bmatrix}.$$

If we make the definition

$$\Pi_1 = GQG^T + A(\Pi_0 - \Pi_0 C^T (C\Pi_0 C^T + R)^{-1} C\Pi_0) A^T$$

and the variable transformation

$$\bar{x} = Ax_0 - GQG^T C^T \lambda_1,$$

then can we obtain the following equivalent representation

$$\begin{bmatrix} \Pi_1^{-1} & C^T & \\ C & & -I \\ & -I & R^{-1} \end{bmatrix} \begin{bmatrix} \bar{x}_0 \\ \lambda_1 \\ x_1 \end{bmatrix} = \begin{bmatrix} \Pi_1^{-1} CAx_0 \\ y_1 \\ 0 \end{bmatrix}$$

The second form is more appealing from a statistical view because it allows us to put our solution in a recursive framework. Using the results from $T = 1$, we obtain the following solution:

$$x_1 = Ax_0 + P_1 C^T (CP_1 C^T + R)^{-1} (y_1 - CAx_0) \tag{D.2}$$

If we define $\hat{x}_1 := Ax_0$ and $\hat{x}_2 := Ax_1$ then we obtain the desired result. It is straightforward to extend the solution for $T > 2$.

Bibliography

Abel, O. and Marquardt, W. (1998), A model predictive control scheme for safe and optimal operation of exothermic semi-batch reactors, *in* 'Proceedings of the 1998 IFAC DYCOPS Symposium, Corfu, Greece', pp. 761–766.

Alamir, M. and Bornard, G. (1995), 'Stability of a truncated infinite constrained receding horizon scheme: The general discrete nonlinear case', *Automatica* **31**, 1352–1356.

Albuquerque, J. and Biegler, L. T. (1996), 'Data reconciliation and gross-error detection for dynamic systems', *AIChE J.* **42**(10), 2841–2856.

Albuquerque, J. and Biegler, L. T. (1997), 'Decomposition algorithms for on-line estimaition with non-linear DAE models', *Comput. Chem. Eng.* **21**(3), 283–299.

Allwright, J. C. and Papavasiliou, G. C. (1992), 'On linear programming and robust model predictive control using impulse-responses', *Sys. Cont. Let.* **18**, 159–164.

Anderson, B. D. O. (1999), 'From Wiener to hidden Markov models', *IEEE Ctl. Sys. Mag.* pp. 41–51.

Arnold, E., Tatjewski, P. and Wolochowicz, P. (1994), 'Two methods for large-scale nonlinear optimization and their comparison on a case study of hydropower optimization', *J. Optim. Theory Appl.* **81**(2), 221–248.

Arrow, K. J. (1951), 'Alternative approaches to the theory of choice in risk-taking situations', *Econometrica* **19**(4), 404–437.

Ascher, U. N. and Petzold, L. R. (1998), *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, SIAM, Philadelphia.

Åström, K. J. (1970), *Introduction to Stochastic Control Theory*, Academic Press, San Diego, California.

Başar, T. (1991), Generalized Riccati equation in dynamic games, *in* S. Bittanti, A. J. Laub and J. C. Willems, eds, 'The Riccati Equation', Springer-Verlag, chapter 11, pp. 293–333.

Başar, T. and Bernhard, P. (1995), *H^{∞}-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, Birkauser, Boston.

Başar, T. and Olsder, G. J. (1999), *Dynamic Noncooperative Game Theory*, SIAM, Philadelphia.

Badgwell, T. A. (1997), 'Robust model predictive control of stable linear systems', *Int. J. Control* **68**(4), 797–818.

Bemporad, A., Mignone, D. and Morari, M. (1999), Moving horizon estimation for hybrid systems and fault detection, *in* 'Proceedings of 1999 American Control Conference, San Diego, California'.

Bemporad, A. and Morari, M. (1999), 'Control of systems integrating logic, dynamics, and constraints', *Automatica* **35**, 407–427.

Bemporad, A. and Mosca, E. (1998), 'Fulfilling hard constrains in uncertain linear systems by reference managing', *Automatica* **34**(3), 451–461.

Bequette, B. W. (1991), 'Nonlinear predictive control using multi-rate sampling', *Can. J. Chem. Eng.* **69**, 136–143.

Berge, C. (1963), *Topological Spaces, including a treatment of Multi-Valued Functions, Vector Spaces and Convexity*, Oliver & Boyd, Edinburgh and London.

Bernhard, P. (1995), 'A discrete-time min-max certainty equivalent principle', *Sys. Cont. Let.* **24**, 229–234.

Bersekas, D. P. (1995*a*), *Dynamic Programming and Optimal Control*, Vol. 1, Athena Scientific, Belmont, MA.

Bersekas, D. P. (1995*b*), *Dynamic Programming and Optimal Control*, Vol. 2, Athena Scientific, Belmont, MA.

Bertsekas, D. P. (1972), On the solution of some minimax problems, *in* 'Proceedings 1972 IEEE Decision and Control Conference', New Orleans, LA, USA, pp. 328–332.

Bertsekas, D. P. (1987), *Dynamic Programming*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Bestle, D. and Zeitz, M. (1983), 'Canonical form observer design for non-linear time-variable systems', *Int. J. Control* **38**(2), 419–431.

Biegler, L. T. (1997), Advances in nonlinear programming concepts for process control, *in* 'IFAC Adchem '97: International Symposium on Advanced Control of Chemical Processes, Banff, Alberta, Canada', pp. 587–598.

Biegler, L. T. (1998), Efficient solution of dynamic optimization and NMPC problems, *in* 'International Symposium on Nonlinear Model Predictive Control, Ascona, Switzerland'.

Binder, T., Blank, L., Dahmen, W. and Marquardt, W. (1998), Towards multiscale dynamic data reconciliation, *in* 'NATO ASI on Nonlinear Model Based Process Control', Kluwer.

Binder, T., Blank, L., Dahmen, W. and Marquardt, W. (1999), 'Regularization of dynamic data reconciliation problems by projection', Submitted for publication in ADCHEM.

Bitmead, R. R., Gevers, M. R., Petersen, I. R. and Kaye, R. J. (1985), 'Monotonicity and stabilizability properties of solutions of the Riccati difference equation: Propositions, lemmas, theorems, fallacious conjectures and counterexamples', *Sys. Cont. Let.* **5**, 309–315.

Bitmead, R. R., Gevers, M. and Wertz, V. (1990), *Adaptive Optimal Control, The Thinking Man's GPC*, Prentice-Hall, Englewood Cliffs, New Jersey.

Bock, H. G., Diehl, M., Leineweber, D. B. and Schlöser, J. P. (1998), A direct multiple shooting method for real-time optimization of nonlinear DAE processes, *in* 'International Symposium on Nonlinear Model Predictive Control, Ascona, Switzerland'.

Braatz, R. D., Tyler, M. L., Morari, M., Pranckh, F. R. and Sartor, L. (1992), 'Identification and cross-directional control of coating processes', *AIChE J.* **38**(9), 1329–1339.

Branicky, M. S., Borkar, V. S. and Mitter, S. K. (1998), 'A unified framework for hybrid control: model and optimal control theory', *IEEE Trans. Auto. Cont.* **43**(1), 31–45.

Bryson, A. E. and Frazier, M. (1963), 'Smoothing for linear and nonlinear dynamic systems', Proc. Optimum Systems Synthesis Conf., U.S. Air Force Tech. Rept. ASD-TDR-063-119.

Bryson, A. E. and Ho, Y. (1975), *Applied Optimal Control*, Hemisphere Publishing, New York.

Callier, F. M. and Desoer, C. A. (1991), *Linear System Theory*, Springer-Verlag, New York.

Camacho, E. and Bordons, C. (1998), *Model Predictive Control*, Springer, Berlin.

Campbell, J. C. and Rawlings, J. B. (1996), Identification and control of sheet forming processes, *in* 'Proceedings of the 13th IFAC World Congress', San Francisco, California.

Campo, P. J. and Morari, M. (1986), ∞-norm formulation of model predicitve control problems, *in* 'Proceedings of the 1986 American Control Conference', pp. 339–343.

Campo, P. J. and Morari, M. (1987), Robust model predictive control, *in* 'Proceedings of the 1987 American Control Conference', pp. 1021–1026.

Campo, P. J. and Morari, M. (1989), 'Model predictive optimal averaging level control', *AIChE J.* **35**(4), 579–591.

Chang, T. S. and Seborg, D. E. (1983), 'A linear programming approach for multivariable feedback control with inequality constraints', *Int. J. Control* **37**, 583–597.

Chen, C. C. and Shaw, L. (1982), 'On receding horizon control', *Automatica* **16**(3), 349–352.

Chen, C.-T. (1984), *Linear System Theory and Design*, Holt, Rhinehart and Winston, New York.

Chen, H. and Allgöwer, F. (1998), 'A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability', *Automatica* **34**(10), 1205–1217.

Chen, H., Scherer, C. W. and Allgöwer, F. (1997), A game theoretic approach to nonlinear robust receding horizon control of constrained systems, *in* 'Proceedings of American Control Conference, Albuquerque, NM', pp. 3073–3077.

Chen, H., Scherer, C. W. and Allgöwer, F. (1998), A robust model predictive control scheme for constrained linear systems, *in* 'Proceedings of the 1998 IFAC DYCOPS Symposium, Corfu, Greece'.

Chisci, L., Lombardi, A. and Mosca, E. (1996), 'Dual-receding horizon control of constrained discrete time systems', *European Journal of Control* **2**, 278–285.

Chmielewski, D. and Manousiouthakis, V. (1996), 'On constrained infinite-time linear quadratic optimal control', *Sys. Cont. Let.* **29**, 121–129.

Congalidis, J. B., Richards, J. R. and Ray, W. H. (1986), Modeling and control of a copolymerization reactor, *in* 'Proceedings of the American Control Conference', Seattle, Washington, pp. 1779–1793.

Cox, H. (1964), 'On the estimation of state variables and parameters for noisy dynamic systems', *IEEE Trans. Auto. Cont.* **9**(1), 5–12.

Cutler, C. R. and Ramaker, B. L. (1980), Dynamic matrix control—a computer control algorithm, *in* 'Proceedings of the Joint Automatic Control Conference'.

Dave, P., Willig, D. A., Kudva, G. K., Pekny, J. F. and Doyle, F. J. (1997), 'LP methods in MPC of large–scale systems: Application to paper–machine CD control', *AIChE J.* **43**(4), 1016–1031.

Davison, E. J. (1973*a*), 'The feedforward control of linear multivariable time-invariant systems', *Automatica* **9**, 561–573.

Davison, E. J. (1973*b*), 'The systematic design of control systems for large multivariable linear time-invariant systems', *Automatica* **9**, 441–452.

De Nicolao, G., Magni, L. and Scattolini, R. (1996), 'On the robustness of receding-horizon control with terminal constraints', *IEEE Trans. Auto. Cont.* **41**(3), 451–453.

De Nicolao, G., Magni, L. and Scattolini, R. (1997), 'Stabilizing receding horizon control of nonlinear ARX models', *Automatica* **33**(9), 169–1697.

De Nicolao, G., Magni, L. and Scattolini, R. (1998), Stability and robustness of nonlinear receding horizon control, *in* 'International Symposium on Nonlinear Model Predictive Control, Ascona, Switzerland'.

de Oliveira, N. M. C. and Biegler, L. T. (1994), 'Constraint handling and stability properties of model-predictive control', *AIChE J.* **40**(7), 1138–1155.

de Souza, C. E., Gevers, M. R. and Goodwin, G. C. (1986), 'Riccati equation in optimal filtering of nonstabilizable systems having singular state transition matrices', *IEEE Trans. Auto. Cont.* **31**(9), 831–838.

Delfour, M. C. and Mitter, S. K. (1969), 'Reachability of perturbed systems and min sup problems', *SIAM J. Cont.* **7**(4), 521–533.

Desborough, L. and Harris, T. (1992), 'Performance assessment measures for univariate feedback control', *Can. J. Chem. Eng.* **70**, 1186–1197.

DeVallière, P. and Bonvin, D. (1990), 'Application of estimation techniques to batch reactors–III. Modeling refinements which improve the quality of state and parameter estimation', *Comput. Chem. Eng.* **14**(7), 799–808.

Doyle, J. C., Glover, K., Khargonekar, P. P. and Francis, B. A. (1989), 'State-space solutions to standard H_2 and H_∞ control problems', *IEEE Trans. Auto. Cont.* **34**(4), 831–847.

Doyle, J. C. and Stein, G. (1979), 'Robustness and observers', *IEEE Trans. Auto. Cont.* **24**(4), 607–611.

Dreyfus, S. E. (1962), 'Some types of optimal control of stochastic systems', *SIAM J. Cont.* **2**(1), 120–134.

Featherstone, A. P. and Braatz, R. D. (1995), Control relevant identification of sheet and film processes, *in* 'Proceedings of the 1995 American Control Conference', Seattle, Washington, pp. 2692–2696.

Findeisen, P. (1997), Moving horizon state estimation of discrete time systems, Master's thesis, University of Wisconsin–Madison.

Fletcher, R. (1987), *Practical Methods of Optimization*, John Wiley & Sons, New York.

Frank, M. and Wolfe, P. (1956), 'An algorithm for quadratic programming', *Naval Res. Logist. Quart.* **3**, 95–110.

García, C. E., Prett, D. M. and Morari, M. (1989), 'Model predictive control: Theory and practice—a survey', *Automatica* **25**(3), 335–348.

Gatzke, E. P. and Doyle III, F. J. (1999), Moving horizon estimation and control of an experimental process, *in* 'AIChE Annual Meeting, Dallas, Texas'.

Gauthier, J. P. and Bornard, G. (1983), 'Commande multivariable en présence de constraintes de type inégalité', *Revue d'Automatique d'Informatique et de Recherche Opérationnelle (RAIRO)* **17**(3), 205–222.

Genceli, H. and Nikolaou, M. (1993), 'Robust stability analysis of constrained l_1–norm model predictive control', *AIChE J.* **39**(12), 1954–1965.

Gesthuisen, R. and Engell, S. (1998), Determination of the mass transport in the polycondensation of polyethyleneterephthalate by nonlinear estimation techniques, *in* 'Proceedings of the 1998 IFAC DYCOPS Symposium, Corfu, Greece'.

Gilbert, E. G. and Tan, K. T. (1991), 'Linear systems with state and control constraints: The theory and application of maximal output admissible sets', *IEEE Trans. Auto. Cont.* **36**(9), 1008–1020.

Gill, P. E., Murray, W., Saunders, M. A. and Wright, M. H. (1983), User's guide for SOL/QPSOL: A Fortran package for quadratic programming, technical report SOL 83-12, Technical report, Systems Optimization Laboratory, Department of Operations Research, Stanford University.

Gill, P. E., Murray, W., Saunders, M. A. and Wright, M. H. (1986), User's guide for SOL/NPSOL (Version 4.0): A Fortran package for nonlinear programming, technical report SOL 86-2, Technical report, Systems Optimization Laboratory, Department of Operations Research, Stanford University.

Glad, T. and Jonson, H. (1984), A method for state and control constrained linear quadratic control problems, *in* 'Proceedings of the 9th IFAC World Congress', Budapest, Hungary.

Golub, G. H. and Van Loan, C. F. (1983), *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Maryland.

Hager, W. W. (1979), 'Lipschitz continuity for constrained processes', *SIAM J. Cont. Opt.* **17**(3), 321–338.

Harris, T. J. (1989), 'Assessment of control loop performance', *Can. J. Chem. Eng.* **67**, 856–861.

Isidori, A. (1989), *Nonlinear Control Systems*, 2nd edn, Springer-Verlag, Berlin.

Jadbabaie, A., Yu, J. and Hauser, J. (1999), 'Unconstrained receding horizon control of nonlinear systems', *IEEE Trans. Auto. Cont.* . Submitted.

James, M. R. and Baras, J. S. (1995), 'Robust H_∞ output feedback control for nonlinear systems', *IEEE Trans. Auto. Cont.* **40**(6), 1007–1017.

Jang, S.-S., Joseph, B. and Mukai, H. (1986), 'Comparison of two approaches to on-line parameter and state estimation of nonlinear systems', *Ind. Eng. Chem. Proc. Des. Dev.* **25**, 809–814.

Jazwinski, A. H. (1970), *Stochastic Processes and Filtering Theory*, Academic Press, New York.

Kailath, T. (1974), 'A view of three decades of linear filtering theory', *IEEE Trans. Inform. Theory* **IT-20**(2), 146–181.

Kalman, R. (1994), 'Randomness reexamined', *Modeling, Identification, and Control* **13**(3), 141–151.

Kalman, R. E. (1960*a*), 'Contributions to the theory of optimal control', *Bull. Soc. Math. Mex.* **5**, 102–119.

Kalman, R. E. (1960*b*), 'A new approach to linear filtering and prediction problems', *Trans. ASME, J. Basic Engineering* pp. 35–45.

Kalman, R. E. and Bucy, R. S. (1961), 'New results in linear filtering and prediction theory', *Trans. ASME, J. Basic Engineering* pp. 95–108.

Keerthi, S. S. (1986), Optimal Feedback Control of Discrete-Time Systems with State-Control Constraints and General Cost Functions, PhD thesis, University of Michigan.

Keerthi, S. S. and Gilbert, E. G. (1985), 'An existence theorem for discrete-time infinite-horizon optimal control problems', *IEEE Trans. Auto. Cont.* **30**(9), 907–909.

Keerthi, S. S. and Gilbert, E. G. (1986), 'Optimal infinite–horizon control and the stabilization of linear discrete–time systems: State–control constraints and nonquadratic cost functions', *IEEE Trans. Auto. Cont.* **31**(3), 264–266.

Keerthi, S. S. and Gilbert, E. G. (1988), 'Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: Stability and moving-horizon approximations', *J. Optim. Theory Appl.* **57**(2), 265–293.

Kim, I., Liebman, M. and Edgar, T. (1991), 'A sequential error-in-variables method for nonlinear dynamic systems', *Comput. Chem. Eng.* **15**(9), 663–670.

Kleinman, D. L. (1970), 'An easy way to stabilize a linear constant system', *IEEE Trans. Auto. Cont.* **15**(12), 692.

Kloeden, P. E. and Platen, E. (1992), *Numerical solution of stochastic differential equations*, Springer Verlag, Berlin.

Kothare, M. V., Balakrishnan, V. and Morari, M. (1996), 'Robust constrained model predictive control using linear matrix inequalities', *Automatica* **32**(10), 1361–1379.

Kozub, D. J. (1997), Controller performance monitoring and diagnosis: Experiences and challenges, *in* J. C. Kantor, C. E. García and B. Carnahan, eds, 'Proceedings of Chemical Process Control – V', CACHE, AIChE, pp. 83–96.

Krener, A. J. and Isidori, A. (1983), 'Linearization by output injection and nonlinear observers', *Sys. Cont. Let.* **3**(1), 47–51.

Kristinsson, K. and Dumont, G. A. (1996), 'Cross–directionalal control on paper machines using Gram polynomials', *Automatica* **32**, 533–548.

Kushner, H. J. (1964), 'On the differential equations satisfied by conditional densities of Markov processes', *SIAM J. Cont. Opt.* **2**, 106.

Kushner, H. J. (1967), 'Dynamical equations for optimal nonlinear filtering', *J. Diff. Eq.* **3**, 179–190.

Kwon, W. H. (1994), Advances in predictive control: theory and application, *in* 'Proceedings of first Asian Control Conference', Tokyo.

Kwon, W. H. (1996), Advances in predictive control: Theory and applications, *in* '1996 Workshop on Model Predictiv Control in Seoul', pp. 1–43.

Kwon, W. H., Bruckstein, A. M. and Kailath, T. (1983), 'Stabilizing state-feedback design via the moving horizon method', *Int. J. Control* **37**(3), 631–643.

Kwon, W. H. and Pearson, A. E. (1977), 'A modified quadratic cost problem and feedback stabilization of a linear system', *IEEE Trans. Auto. Cont.* **22**(5), 838–842.

Kwon, W. H. and Pearson, A. E. (1978), 'On feedback stabilization of time-varying discrete linear systems', *IEEE Trans. Auto. Cont.* **23**(3), 479–481.

Lee, J. H. and Cooley, B. (1997), Recent advances in model predictive control and other related areas, *in* J. C. Kantor, C. E. García and B. Carnahan, eds, 'Proceedings of Chemical Process Control – V', CACHE, AIChE, pp. 201–216.

Lee, J. H. and Yu, Z. (1997), 'Worst-case formulations of model predictive control for systems with bounded parameters', *Automatica* **33**(5), 763–781.

Liebman, M., Edgar, T. and Lasdon, L. (1992), 'Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques', *Comput. Chem. Eng.* **16**(10/11), 963–986.

Lim, A., Moore, J. and Faybusovich, L. (1996), Linearly constrained LQ and LQG optimal control, *in* 'Proceedings of the 13th IFAC World Congress', San Francisco, California.

Lin, W. and Byrnes, C. I. (1996), 'H_∞-control of discrete-time nonlinear systems', *IEEE Trans. Auto. Cont.* **41**(4), 494–510.

Ljung, L. (1979), 'Asymptotic behaviour of the extended Kalman filter as a parameter estimator for linear systems', *IEEE Trans. Auto. Cont.* **24**(1), 36–50.

Luenberger, D. G. (1966), 'Observers for multivariable systems', *IEEE Trans. Auto. Cont.* **11**(2), 190–197.

Luenberger, D. G. (1969), *Optimization by Vector Space Methods*, John Wiley and Sons, Inc., New York.

Luenberger, D. G. (1971), 'An introduction to observers', *IEEE Trans. Auto. Cont.* **16**(2), 596–602.

Luenberger, D. G. (1992), 'A double look at duality', *IEEE Trans. Auto. Cont.* **37**(10), 1474–1482.

Magni, L. (1998), Nonlinear Receding Horizon Control: Theory and Application, PhD thesis, University of Pavia.

Magni, L. and Sepulchre, R. (1997), 'Stability margins of nonlinear receding-horizon control via inverse optimality', *Sys. Cont. Let.* **32**, 241–245.

Mangasarian, O. (1994), *Nonlinear Programming*, SIAM, Philadelphia, PA.

Mayne, D. Q. (1995*a*), Optimization in model based control, *in* 'Proceedings of the IFAC Symposium Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Processes', Helsingor, Denmark, pp. 229–242.

Mayne, D. Q. (1995*b*), Optimization in model predictive control, *in* R. Berber, ed., 'Methods of Model Based Process Control', Nato Advanced Study Institute series: E Applied Sciences 293, Kluwer, Dordrecht, The Netherlands, pp. 367–398.

Mayne, D. Q. (1997), Nonlinear model predictive control: An assessment, *in* J. C. Kantor, C. E. García and B. Carnahan, eds, 'Proceedings of Chemical Process Control – V', CACHE, AIChE, pp. 217–231.

Mayne, D. Q. and Michalska, H. (1990), 'Receding horizon control of nonlinear systems', *IEEE Trans. Auto. Cont.* **35**(7), 814–824.

Mayne, D. Q., Rawlings, J. B., Rao, C. V. and Scokaert, P. O. M. (1999), 'Model predictive control: a review', Accepted for publication in *Automatica*.

Mayne, D. Q., Rawlings, J. B., Rao, C. V. and Scokaert, P. O. M. (2000), 'Constrained model predictive control: Stability and optimality'. Automatica. In press.

Meadows, E. S., Henson, M. A., Eaton, J. W. and Rawlings, J. B. (1995), 'Receding horizon control and discontinuous state feedback stabilization', *Int. J. Control* **62**(5), 1217–1229.

Mehrotra, S. (1992), 'On the implementation of a primal-dual interior point method', *SIAM J. Optim.* **2**(1), 575–601.

M'hamdi, A., Helbig, A., Abel, O. and Marquardt, W. (1996), Newton-type receding horizon control and state estimation, *in* 'Proceedings of the 1996 IFAC World Congress', San Francisco, California, pp. 121–126.

Michalska, H. and Mayne, D. Q. (1991), 'Receding horizon control of nonlinear systems without differentiability of the optimal value function', *Sys. Cont. Let.* **16**, 123–130.

Michalska, H. and Mayne, D. Q. (1993), 'Robust receding horizon control of constrained nonlinear systems', *IEEE Trans. Auto. Cont.* **38**(11), 1623–1633.

Michalska, H. and Mayne, D. Q. (1995), 'Moving horizon observers and observer-based control', *IEEE Trans. Auto. Cont.* **40**(6), 995–1006.

Moraal, P. E. and Grizzle, J. W. (1995), 'Observer design for nonlinear systems with discrete-time measurements', *IEEE Trans. Auto. Cont.* **40**(3), 395–404.

Morari, M. and Lee, J. H. (1997), Model predictive control: past, present and future, *in* 'Proceedings of joint 6th international symposium on process systems engineering (PSE '97) and 30th European symposium on computer aided process systems engineering (ESCAPE 7)', Trondheim, Norway.

Morshedi, A. M., Cutler, C. R. and Skrovanek, T. A. (1985), Optimal solution of dynamic matrix control with linear programming techniques (LDMC), *in* 'Proceedings of the 1985 American Control Conference', pp. 199–208.

Muske, K. R. (1997), Steady–state target optimizaton in linear model predictive control, *in* 'Proceedings of the American Control Conference', Albuquerque, NM, pp. 3597–3601.

Muske, K. R. and Edgar, T. F. (1996), Nonlinear state estimation, *in* M. A. Henson and D. E. Seborg, eds, 'Nonlinear Process Control', Prentice Hall, pp. 311–370.

Muske, K. R. and Rawlings, J. B. (1993), 'Model predictive control with linear models', *AIChE J.* **39**(2), 262–287.

Muske, K. R. and Rawlings, J. B. (1995), Nonlinear moving horizon state estimation, *in* R. Berber, ed., 'Methods of Model Based Process Control', Nato Advanced Study Institute series: E Applied Sciences 293, Kluwer, Dordrecht, The Netherlands, pp. 349–365.

Muske, K. R., Rawlings, J. B. and Lee, J. H. (1993), Receding horizon recursive state estimation, *in* 'Proceedings of the 1993 American Control Conference', pp. 900–904.

Nocedal, J. and Wright, S. J. (1999), *Numerical Optimization*, Springer-Verlag, New York.

Ogunnaike, B. A. (1995), A contemporary industrial perspective on process control theory and practice, *in* 'Proceedings of the IFAC Symposium Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Processes', Helsingor, Denmark, pp. 1–8.

Outraka, J. V. (1976), 'Discrete optimal control problems with nonsmooth costs', *Kybernetika* **12**(3), 192–205.

Panier, E. R. and Tits, A. L. (1993), 'On combining feasibility, descent and superlinear convergence in inequality constrained optimization', *Mathematical Programming* **59**, 261–276.

Parisini, T. and Zoppoli, R. (1995), 'A receding-horizon regulator for nonlinear systems and a neural approximation', *Automatica* **31**(10), 1443–1451.

Polak, E. (1997), *Optimization: Algorithms and Consistent Approximations*, Springer Verlag, New York. ISBN 0-387-94971-2.

Polak, E. and Yang, T. H. (1993*a*), 'Moving horizon control of linear systems with input saturation and plant uncertainty–part 1: Robustness', *Int. J. Control* **58**(3), 613–638.

Polak, E. and Yang, T. H. (1993*b*), 'Moving horizon control of linear systems with input saturation and plant uncertainty–part 2: Disturbance rejection and tracking', *Int. J. Control* **58**(3), 639–663.

Prett, D. M. and Morari, M., eds (1987), *The Shell Process Control Workshop*, Butterworths, Boston.

Propoi, A. I. (1963), 'Use of linear programming methods for synthesizing sampled-data automatic systems', *Automn. Remote Control* **24**(7), 837–844.

Qin, J. S. and Badgwell, T. A. (1998), An overview of nonlinear model predictive control applications, *in* 'International Symposium on Nonlinear Model Predictive Control, Ascona, Switzerland'.

Qin, S. J. (1998), 'Controller performance monitoring – a review and assessment', *Comput. Chem. Eng.* **23**, 173–186.

Qin, S. J. and Badgwell, T. A. (1997), An overview of industrial model predictive control technology, *in* J. C. Kantor, C. E. García and B. Carnahan, eds, 'Proceedings of Chemical Process Control–V', CACHE, AIChE, pp. 232–256.

Ramamurthi, Y., Sistu, P. and Bequette, B. (1993), 'Control-relevant dynamic data reconciliation and parameter estimation', *Comput. Chem. Eng.* **17**(1), 41–59.

Rao, C. V., Campbell, J. C., Rawlings, J. B. and Wright, S. J. (1997), Efficient implementation of model predictive control for sheet and film forming processes, *in* 'Proceedings of American Control Conference, Albuquerque, NM', pp. 2940–2944.

Rao, C. V. and Rawlings, J. B. (1998*a*), Nonlinear moving horizon estimation, *in* 'International Symposium on Nonlinear Model Predictive Control, Ascona, Switzerland'.

Rao, C. V. and Rawlings, J. B. (1998*b*), Optimization strategies for linear model predictive control, *in* 'Proceedings of the 1998 IFAC DYCOPS Symposium, Corfu, Greece'.

Rao, C. V. and Rawlings, J. B. (1999), 'Steady states and constraints in model predictive control', *AIChE J.* **45**(6), 1266–1278.

Rao, C. V. and Rawlings, J. B. (2000), 'Linear programming and model predictive control', *J. Proc. Cont.* **10**, 283–289.

Rao, C. V., Rawlings, J. B. and Lee, J. H. (1999), 'Constrained linear state estimation - a moving horizon approach', Submitted for publication in *Automatica*.

Rao, C. V., Rawlings, J. B. and Mayne, D. Q. (2000), 'Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations', Submitted for publication in *IEEE Transactions on Automatic Control*.

Rao, C. V., Wright, S. J. and Rawlings, J. B. (1998), 'On the application of interior point methods to model predictive control', *J. Optim. Theory Appl.* **99**, 723–757.

Rauch, H. E., Tung, F. and Striebel, C. T. (1965), 'Maximum likelihood estimates of linear dynamic systems', *AIAA J.* **3**(8), 1445–1450.

Rawlings, J. B. (1999), Tutorial: Model predictive control technology, *in* 'Proceedings of the American Control Conference, San Diego, CA', pp. 662–676.

Rawlings, J. B., Jerome, N. F., Hamer, J. W. and Bruemmer, T. M. (1989), End-point control in semi-batch chemical reactors, *in* 'Proceedings of the IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns, and Batch Processes', pp. 323–328.

Rawlings, J. B., Meadows, E. S. and Muske, K. R. (1994), Nonlinear model predictive control: a tutorial and survey, *in* 'ADCHEM '94 Proceedings, Kyoto, Japan', pp. 185–197.

Rawlings, J. B. and Muske, K. R. (1993), 'Stability of constrained receding horizon control', *IEEE Trans. Auto. Cont.* **38**(10), 1512–1516.

Ray, W. H. and Barney, J. R. (1972), 'The application of differential game theory to process control problems', *Chem. Eng. J.* **3**, 237–244.

Richalet, J., Rault, A., Testud, J. L. and Papon, J. (1978), 'Model predictive heuristic control: Applications to industrial processes', *Automatica* **14**, 413–428.

Ricker, N. L., Subrahmanian, T. and Sim, T. (1988), Case studies of model-predictive control in pulp and paper production, *in* T. J. McAvoy, Y. Arkun and E. Zafiriou, eds, 'Proceedings of the 1988 IFAC Workshop on Model Based Process Control', Pergamon Press, Oxford, pp. 13–22.

Robertson, D. (1996), Development and Statistical Intepretation of Tools for Nonlinear Estimation, PhD thesis, Auburn University.

Robertson, D. G. and Lee, J. H. (1995), 'A least squares formulation for state estimation', *J. Proc. Cont.* **5**(4), 291–299.

Robertson, D. G. and Lee, J. H. (1998), 'On the use of constraints in least squares estimation and control', Submitted for Publication.

Robertson, D. G., Lee, J. H. and Rawlings, J. B. (1996), 'A moving horizon-based approach for least-squares state estimation', *AIChE J.* **42**(8), 2209–2224.

Rockafellar, R. T. (1970), *Convex Analysis*, Princeton University Press, Princeton, N.J.

Rossiter, J. A., Rice, M. J. and Kouvaritakis, B. (1997), A robust stable state-space approach to stable predictive control stategies, *in* 'Proceedings of the American Control Conference', Albuquerque, NM, pp. 1640–1641.

Royden, H. L. (1988), *Real Analysis*, 3rd edn, Prentice Hall, Englewood Cliffs, New Jersey.

Russo, L. P. and Young, R. E. (1999), Moving horizon state estimation applied to an industrial polymerization process, *in* 'Proceedings of 1999 American Control Conference, San Diego, California'.

Safonov, M. G. and Athans, M. (1978), 'Robustness and computational aspects of nonlinear stochastic estimators and regulators', *IEEE Trans. Auto. Cont.* **23**(4), 717–725.

Santos, L. O. and Bieger, L. T. (1999), 'A tool to analyze robust stability for model predictive control', *J. Proc. Cont.* **9**, 233–245.

Schuler, H. and Schmidt, C.-U. (1992), 'Calorimetric-state estimators for chemical reactor diagnosis and control: review of methods and applications', *Chem. Eng. Sci.* **47**(4), 899–915.

Scokaert, P. O. M., Mayne, D. Q. and Rawlings, J. B. (1999), 'Suboptimal model predictive control (feasibility implies stability)', *IEEE Trans. Auto. Cont.* **44**(3), 648–654.

Scokaert, P. O. M. and Rawlings, J. B. (1996), Infinite horizon linear quadratic control with constraints, *in* 'Proceedings of the 13th IFAC World Congress', San Francisco, California.

Scokaert, P. O. and Mayne, D. Q. (1998), 'Min–max feedback model predictive control for constrained linear systems', *IEEE Trans. Auto. Cont.* **43**(8), 1136–1142.

Scokaert, P. O. and Rawlings, J. B. (1998), 'Constrained linear quadratic regulation', *IEEE Trans. Auto. Cont.* **43**(8), 1163–1169.

Scokaert, P. O. and Rawlings, J. B. (1999), 'Feasibility issues in linear model predictive control', *AIChE J.* **45**(8), 1649–1659.

Scokaert, P. O., Rawlings, J. B. and Meadows, E. S. (1997), 'Discrete-time stability with perturbations: Application to model predictive control', *Automatica* **33**(3), 463–470.

Shamma, J. S. and Xiong, D. (1997), 'Linear nonquadratic optimal control', *IEEE Trans. Auto. Cont.* **42**(6), 875–879.

Skogestad, S. and Postlethwaite, I. (1996), *Multivariable Feedback Control, Analysis and Design*, John Wiley & Sons, New York, N.Y.

Slotine, J. J. E., Hedrick, J. K. and Misawa, E. A. (1987), 'On sliding observers for nonlinear systems', *J. Dynam. Sys. Meas. Cont.* **109**(3), 245–252.

Slupphaug, O. and Foss, B. A. (1997), Model predictive control for a class of hybrid systems, *in* 'Proceedings of European Control Conference, Brussels, Belgium'.

Song, Y. and Grizzle, J. W. (1995), 'The extended Kalman filter as a local asymptotic observer for discrete-time nonlinear systems', *J. Math. Sys., Estim, and Control.* **5**(1), 59–78.

Sontag, E. D. (1990), *Mathematical Control Theory*, Springer-Verlag, New York.

Sorenson, H. W. (1970), 'Least–squares estimation: from Gauss to Kalman', *IEEE Spectrum* pp. 63–68.

Steinbach, M. C. (1994), A structured interior point SQP method for nonlinear optimal control problems, *in* R. Burlirsch and D. Kraft, eds, 'Computational Optimal Control', Birkhäuser, pp. 213–222.

Stephens, C. P. and Baritompa, W. (1998), 'Global optimization requires global information', *J. Optim. Theory Appl.* **96**, 575–588.

Stratonovich, R. L. (1960), 'Conditional Markov processes', *Theor. Probability Appl.* **5**, 156–178.

Striebel, C. (1965), 'Sufficient statistics in the optimum control of stochastic systems', *J. Math. Anal. Appl.* **12**, 576–592.

Sznaier, M. and Damborg, M. J. (1987), Suboptimal control of linear systems with state and control inequality constraints, *in* 'Proceedings of the 26th Conference on Decision and Control', Los Angeles, CA, pp. 761–762.

Thomas, Y. A. (1975), 'Linear quadratic optimal estimation and control with receding horizon', *Electron. Lett.* **11**, 19–21.

Tjoa, I. B. and Biegler, L. T. (1991), 'Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems', *Comput. Chem. Eng.* **15**(10), 679–690.

Tyler, M. L. (1997), Performance Monitoring and Fault Detection in Control Systems, PhD thesis, California Institute of Technology.

Tyler, M. L. and Morari, M. (1996), 'Stability of constrained moving horizon estimation schemes', *Preprint AUT96–18*, Automatic Control Laboratory, Swiss Federal Institute of Technology.

Uppal, A., Ray, W. H. and Poore, A. B. (1974), 'On the dynamic behavior of continuous stirred tank reactors', *Chem. Eng. Sci.* **29**, 967–985.

van der Schaft, A. (1996), *L_2 Gain and Passivity Techniques in Nonlinear Control*, Springer Verlag, Berlin.

Verdu, S. and Poor, H. V. (1987), 'Abstract dynamic programming models under commutativity conditions', *SIAM J. Cont. Opt.* **25**, 990–1006.

Vidyasagar, M. (1980), 'On the stabilization of nonlinear systems using state detection', *IEEE Trans. Auto. Cont.* **25**, 504–409.

Walcott, B., Corless, M. and Żak, S. (1987), 'Comparative study of non–linear state–observation techniques', *Int. J. Control* **45**(6), 2109–2132.

Wilhelm, Jr., R. G. and Fjeld, M. (1983), Control algorithms for cross directional control: the state of the art, *in* 'Proceedings of the 5th IFAC/IMEKO Conference on Instrumentation in the Paper, Rubber, Plastic and Polymerization Industries', Antwerp.

Willems, J. C. (1972), 'Dissipative dynamic systems part 1: general theory', *Arch. Rational Mech. Anal.* **45**, 321–351.

Wonham, W. M. (1969), 'Optimal stochastic control', *Automatica* **5**, 113–118.

Wonham, W. M. (1985), *Linear Multivariable Control: A Geometric Approach*, third edn, Springer-Verlag, New York.

Wright, S. J. (1993), 'Interior point methods for optimal control of discrete-time systems', *J. Optim. Theory Appl.* **77**, 161–187.

Wright, S. J. (1996), 'Modified Cholesky factorizations in interior-point algorithms for linear programming', *Preprint MCS-P600-0596*, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Ill.

Wright, S. J. (1997*a*), Applying new optimization algorithms to model predictive control, *in* J. C. Kantor, C. E. García and B. Carnahan, eds, 'Chemical Process Control–V', CACHE, AIChE, pp. 147–155.

Wright, S. J. (1997*b*), *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia.

Wright, S. J. (1997*c*), 'Stability of augmented system factorizations in interior-point methods', *SIAM J. Matrix Anal. and Appl.* **18**, 191–222.

Yang, Z. and Polak, E. (1993), 'Moving horizon control of nonlinear systems with input saturation, disturbances and plant uncertainty', *Int. J. Control* **58**, 875–903.

Zadeh, L. A. and Whalen, B. H. (1962), 'On optimal control and linear programming', *IRE Trans. Auto, Cont.* **AC–7**, 45–46.

Zafiriou, E. and Marchal, A. L. (1991), 'Stability of SISO quadratic dynamic matrix control with hard output constraints', *AIChE J.* **37**(10), 1550–1560.

Zheng, A. (1999), 'Robust stability analysis of constrained model predictive control', *J. Proc. Cont.* **9**, 271–278.

Zheng, A. and Morari, M. (1995), 'Stability of model predictive control with mixed constraints', *IEEE Trans. Auto. Cont.* **40**(10), 1818–1823.

Zheng, Z. Q. and Morari, M. (1993), Robust stability of constrained model predictive control, *in* 'Proceedings of the 1993 American Control Conference', pp. 379–383.

Zimmer, G. (1994), 'State observation by on-line minimization', *Int. J. Control* **60**(4), 595–606.

Vita

Christopher Vavilikolanu Rao was born in Huntingdon, Pennsylvania on June 22nd, 1972. At the age of 8, his family moved to New York City. He graduated from Riverdale Country School in 1990, and in the fall of that year attended Carnegie Mellon University. He received his Bachelor of Science degree in Chemical Engineering in May, 1994 from Carnegie Mellon University. He began his graduate studies in the Fall of 1994 when he joined the research group of Dr. James B. Rawlings in the Department of Chemical Engineering at The University of Texas at Austin. In June 1995, he moved with Dr. Rawlings from Austin to the Department of Chemical Engineering at the University of Wisconsin-Madison.

Permanent Address: 220 South Lyndon Lane,
Louisville, KY 40222

This dissertation was typeset by the author using LaTeX 2_ε[1].

[1]This particular University of Wisconsin complient style was modified from The University of Texas at Austin styles as written by Dinesh Das (LaTeX 2_ε), Khe–Sing The (LaTeX), and John Eaton (LaTeX).